

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:37:31 ; Search time 11.7951 Seconds  
(without alignments)  
1180.438 Million cell updates/sec

Title: US-09-762-261C-1\_COPY\_301\_336  
Perfect score: 201  
Sequence: 1 CSRPNNTKSPMGPRAPFTYTGIIIGDROAHCHN 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_16Dec04:\*
- 1: Geneseqp1980s:\*
  - 2: Geneseqp1990s:\*
  - 3: Geneseqp2000s:\*
  - 4: Geneseqp2001s:\*
  - 5: Geneseqp2002s:\*
  - 6: Geneseqp2003as:\*
  - 7: Geneseqp2003bs:\*
  - 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	866	3	AAY79020 HIV-1 env
2	195	97.0	35	3	AAY79023 HIV-1 RV
3	187	93.0	842	3	AAB69350 HIV-1 non
4	185	92.0	35	2	AAR20883 PND EE378
5	185	92.0	35	2	AAR20953 PND EE669
6	185	92.0	35	2	AAR20853 PND EE322
7	183	91.0	35	2	AAR21134 PND EE228
8	183	91.0	35	2	AAR20911 PND EE533
9	183	91.0	35	2	AAR21135 PND EE228
10	183	91.0	35	2	AAR20912 PND EE533
11	183	91.0	35	2	AAR20910 PND EE533
12	183	91.0	35	2	AAR21136 PND EE228
13	183	91.0	35	2	AAR29113 Group I1
14	183	91.0	45	8	ADR58150 Novel ant
15	182	90.5	39	2	AAW84056 HIV-1 Ba-
16	182	90.5	43	2	AAW63735 HIV-1 JR-
17	182	90.5	45	8	ADR58145 Novel ant
18	182	90.5	60	2	AAR68028 HIV-1 gpl
19	182	90.5	376	6	ABU07702 Viral coa
20	182	90.5	473	4	ABU61502 HIV-1 gpl
21	182	90.5	473	6	ABR57050 HIV-1 JR-
22	182	90.5	473	7	ABR82838 HIV env g
23	182	90.5	473	7	ADF18104 HIV-1 str
24	182	90.5	474	8	ADP74743 HIV-1 iso
25	182	90.5	474	8	ADP74742 HIV-1 iso

26	182	90.5	502	4	AAB82953 HIV-1 JR-
27	182	90.5	506	6	ABU07703 Viral coa
28	182	90.5	506	6	ABU07700 Viral coa
29	182	90.5	507	2	AAR60789 HIV virus
30	182	90.5	507	2	AAR60783 HIV virus
31	182	90.5	556	6	ABU07696 Viral coa
32	182	90.5	579	4	ABU61506 HIV-1 del
33	182	90.5	579	6	ABR57053 HIV-1 JR-
34	182	90.5	579	7	ADF18109 HIV-1 JR-
35	182	90.5	579	8	ADK14401 HIV gpl40
36	182	90.5	590	6	ABU07697 Viral coa
37	182	90.5	643	4	AAB61505 HIV-1 SOS
38	182	90.5	643	6	ABR57052 HIV-1 JR-
39	182	90.5	643	7	ABR82839 HIV env g
40	182	90.5	643	7	ADF18107 HIV-1 JR-
41	182	90.5	643	8	ADK14399 HIV gpl40
42	182	90.5	720	6	ABU07694 Viral coa
43	182	90.5	720	6	ABU07695 Viral coa
44	182	90.5	847	8	ADK14404 HIV mutan
45	182	90.5	847	8	ADK14405 HIV mutan

ALIGNMENTS

RESULT 1  
AAY79020  
ID AAY79020 standard; protein; 866 AA.

XX AAY79020;  
XX  
XX  
DT 12-SEP-2003 (revised)  
DT 05-JUN-2000 (first entry)  
XX  
XX HIV-1 envelope protein amino acid sequence.  
XX  
KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;  
KW vaccine; antibody; detect.  
XX

OS Human immunodeficiency virus 1.

XX WO200007631-A1.

XX 17-FEB-2000.

XX 04-AUG-1999; 99WO-US017596.

XX 04-AUG-1998; 98US-0095267P.

XX (JACK-) JACKSON FOUND HENRY M.

XX Quinnan GV, Zhang PF;

XX WPI; 2000-205578/18.

PT Human immunodeficiency virus-1 envelope protein, useful as a vaccine or immunogenic preparation against HIV-1 infection induces the production of broadly cross-reactive neutralizing anti-serum.

XX Claim 2; Page 33; 54pp; English.

XX This sequence represents a human immunodeficiency virus type 1 (HIV-1) envelope (env) protein amino acid sequence. The invention relates to the HIV-1 env protein or its fragments, which when administered to a mammal, induces the production of broadly cross-reactive neutralising anti-serum against multiple strains of HIV-1. The HIV-1 env protein and its fragments are useful as vaccines, immunogenic compositions or diagnostic reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing antibodies in a sample. The amino acid sequence and its fragments or mRNA sequence are useful for generating antibodies in a mammal. In addition, a recombinant delivery vector containing the env amino acid sequence may also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 866 AA;  
 Query Match 100.0%; Score 201; DB 3; Length 866;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-18;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 36  
 |||||  
 Db 301 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 336

RESULT 2  
 AAY79023  
 ID AAY79023 standard; protein; 35 AA.  
 XX AAY79023;

XX 12-SEP-2003 (revised)  
 DT 05-JUN-2000 (first entry)  
 XX HIV-1 RV strain envelope protein V3 region peptide sequence #2.  
 DE HIV-1; envelope protein; env; cross reactive anti-serum; antibody;  
 KW vaccine; antibody; V3 region; detect.  
 XX Human immunodeficiency virus 1.  
 OS  
 XX WO200007631-A1.

XX 17-FEB-2000.  
 XX 04-AUG-1999; 99WO-US017596.  
 XX 04-AUG-1998; 98US-0095267P.  
 PR (JACK-) JACKSON FOUND HENRY M.  
 XX Quinnan GV, Zhang PF;  
 PI WPI; 2000-205578/18.  
 XX Human immunodeficiency virus-1 envelope protein, useful as a vaccine or  
 PT immunogenic preparation against HIV-1 infection induces the production of  
 PT broadly cross-reactive neutralizing anti-serum.  
 XX Example 5; Page 27; 54pp; English.

XX This sequence represents a human immunodeficiency virus type 1 (HIV-1) R2  
 CC strain envelope (env) protein V3 region peptide sequence. The invention  
 CC relates to the HIV-1 env protein or its fragments, which when  
 CC administered to a mammal, induces the production of broadly cross-  
 CC reactive neutralising anti-serum against multiple strains of HIV-1. The  
 CC HIV-1 env protein and its fragments are useful as vaccines, immunogenic  
 CC compositions or diagnostic reagents. They can also be used to detect anti-  
 CC -HIV-1/HIV-1 neutralizing antibodies in a sample. The amino acid sequence  
 CC and its fragments or mRNA sequence are useful for generating antibodies  
 CC in a mammal. In addition, a recombinant delivery vector containing the  
 CC env amino acid sequence may also be used as a vaccine. (Updated on 12-SEP  
 CC -2003 to standardise OS field)

XX SQ Sequence 35 AA;  
 Query Match 97.0%; Score 195; DB 3; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 1e-16;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 |||||  
 Db 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35

RESULT 3

AAB69350  
 ID AAB69350 standard; protein; 842 AA.  
 XX AAB69350;  
 AC 12-SEP-2003 (revised)  
 DT 20-APR-2001 (first entry)  
 XX HIV-1 non-subtype B clone 93BR029-4 env protein.  
 DE HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env; vpr;  
 KW vif; vpr; tat; rev; nef; vaccine.  
 XX Human immunodeficiency virus 1.  
 OS  
 XX WO200026416-A1.

XX 11-MAY-2000.  
 XX 25-OCT-1999; 99WO-US024837.  
 XX 02-NOV-1998; 98US-00184418.  
 PR (UABR-) UAB RES FOUND.  
 PA Hahn BH, Shaw GM, Gao F;  
 PI WPI; 2000-365651/31.

XX Novel genomic nucleic acids of non-subtype B human immunodeficiency virus  
 PT type 1 useful for detecting and treating AIDS comprises a specific  
 PT nucleotide sequence.

XX Claim 41; Fig 21; 131pp; English.

XX The present invention provides the protein and coding sequences for a  
 CC number of human immunodeficiency virus (HIV) type 1 non-subtype B  
 CC isolates. The sequences shown include the near full-length coding  
 CC sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,  
 CC rev and nef proteins. These can be used to detect the presence of HIV-1  
 CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.  
 CC These antibodies can be used in vaccines to prevent and treat HIV  
 CC infection. (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 842 AA;

Query Match 93.0%; Score 187; DB 3; Length 842;  
 Best Local Similarity 88.9%; Pred. No. 3.2e-16;  
 Matches 32; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 36  
 |||||  
 Db 298 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHCN 333

RESULT 4  
 AAR20883  
 ID AAR20883 standard; protein; 35 AA.  
 AC AAR20883;

XX 01-JUN-1992 (first entry)

XX PND EE378-3.  
 XX HIV; PND; principal neutralising determinant; Omp; Neisseria;  
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.  
 XX Human immunodeficiency virus.  
 OS  
 XX EP471407-A.

XX 19-FEB-1992.

```

XX PF 07-AUG-1991; 91EP-00202025.
XX PF 13-AUG-1990; 90US-00566638.
XX PR 13-AUG-1990; 90US-00566654.
XX PR 13-AUG-1990; 90US-00566656.
XX PA (MERI ) MERCK & CO INC.
XX PI Lewis JA, Davide JP, Waterbury JA;
XX XX
XX DR N-PSDB; AAQ21298.
XX XX
XX PT New antigenic conjugate of HIV major neutralisation determinant -
XX PT covalently linked to outer membrane proteosome of Neisseria, useful as
XX PT vaccine and in treating and preventing HIV infections, AIDS and ARC.
XX PS Claim 1; Page 175 + 117-118; 177pp; English.
XX XX
XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
XX CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
XX CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
XX CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
XX CC then linked by cross-linking agents, monogeneric spacers or bigeneric
XX CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
XX CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
XX CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
XX CC post- exposure to prevent or treat HIV infection or disease, and are
XX CC capable of producing an immune response specific for the immunogen. See
XX CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78
XX SQ Sequence 35 AA;
XX
XX Query Match 92.0%; Score 185; DB 2; Length 35;
XX Best Local Similarity 91.4%; Pred. No. 2.3e-17;
XX Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35
XX |:|||||:|||||:|||||:|||||:|||||
XX Db 1 CTRPNNTKRSIPGPGAFYTTGQIIGDIRQAH 35
XX
XX RESULT 6
XX AAR20853
XX ID AAR20853 standard; protein; 35 AA.
XX XX
XX AC AAR20853;
XX XX
XX DT 01-JUN-1992 (first entry)
XX XX
XX DE PND EE322-2.
XX XX
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX XX
XX OS Human immunodeficiency virus.
XX XX
XX PN EP471407-A.
XX XX
XX PD 19-FEB-1992.
XX XX
XX PF 07-AUG-1991; 91EP-00202025.
XX XX
XX PR 13-AUG-1990; 90US-00566638.
XX PR 13-AUG-1990; 90US-00566654.
XX PR 13-AUG-1990; 90US-00566656.
XX XX
XX PA (MERI ) MERCK & CO INC.
XX XX
XX PI Lewis JA, Davide JP, Waterbury JA;
XX XX
XX DR WPI; 1992-058471/08.
XX DR N-PSDB; AAQ21268.
XX XX
XX PT New antigenic conjugate of HIV major neutralisation determinant -
XX PT covalently linked to outer membrane proteosome of Neisseria, useful as
XX PT vaccine and in treating and preventing HIV infections, AIDS and ARC.
XX XX
XX PS Claim 1; Page 175 + 93-94; 177pp; English.
XX XX
XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
XX CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
XX CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
XX CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
XX CC then linked by cross-linking agents, monogeneric spacers or bigeneric
XX CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
XX CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
XX CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
XX CC post- exposure to prevent or treat HIV infection or disease, and are
XX CC capable of producing an immune response specific for the immunogen. See
XX CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78
XX SQ Sequence 35 AA;
XX
XX Query Match 92.0%; Score 185; DB 2; Length 35;
XX Best Local Similarity 91.4%; Pred. No. 2.3e-17;
XX Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35
XX |:|||||:|||||:|||||:|||||:|||||
XX Db 1 CTRPNNTKRSIPGPGAFYTTGQIIGDIRQAH 35
XX
XX RESULT 5
XX AAR20953
XX ID AAR20953 standard; protein; 35 AA.
XX XX
XX AC AAR20953;
XX XX
XX DT 01-JUN-1992 (first entry)
XX XX
XX DE PND EE669-3.
XX XX
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX XX
XX OS Human immunodeficiency virus.
XX XX
XX PN EP471407-A.
XX XX
XX PD 19-FEB-1992.
XX XX
XX PF 07-AUG-1991; 91EP-00202025.
XX XX
XX PR 13-AUG-1990; 90US-00566638.
XX PR 13-AUG-1990; 90US-00566654.
XX PR 13-AUG-1990; 90US-00566656.
XX XX
XX PA (MERI ) MERCK & CO INC.
XX XX
XX PI Lewis JA, Davide JP, Waterbury JA;
XX XX
XX DR WPI; 1992-058471/08.

```

CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
CC post- exposure to prevent or treat HIV infection or disease, and are  
CC capable of producing an immune response specific for the immunogen. See  
CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 92.0%; Score 185; DB 2; Length 35;  
Best Local Similarity 91.4%; Pred. No. 2.3e-17;  
Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 35  
|:|||||||||:|||||||||:|||||||||:|||||  
Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAH 35

RESULT 7

AAR21134  
ID AAR21134 standard; protein; 35 AA.

XX AAR21134;

XX 01-JUN-1992 (first entry)

XX PND EE228-1.

XX HIV; PND; principal neutralising determinant; Omp; Neisseria;  
XX outer membrane proteosome; AIDS; vaccine; envelope gene.

XX Human immunodeficiency virus.

XX EP471407-A.

XX 19-FEB-1992.

XX 07-AUG-1991; 91EP-00202025.

XX 13-AUG-1990; 90US-00566638.

XX 13-AUG-1990; 90US-00566654.

XX 13-AUG-1990; 90US-00566656.

XX (MERI ) MERCK & CO INC.

XX Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21232.

XX New antigenic conjugate of HIV major neutralisation determinant -

XX covalently linked to outer membrane proteosome of Neisseria, useful as

XX vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 64; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
XX encode PNDs of HIV (envelope protein) which are useful as immunogens for  
XX AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
XX conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,  
XX then linked by cross-linking agents, monogeneric spacers or bigeneric  
XX spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
XX or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
XX Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
XX post- exposure to prevent or treat HIV infection or disease, and are  
XX capable of producing an immune response specific for the immunogen. See  
XX also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 91.0%; Score 183; DB 2; Length 35;  
Best Local Similarity 91.4%; Pred. No. 4.2e-17;

Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 35  
|:|||||||||:|||||||||:|||||||||:|||||  
Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAH 35

RESULT 8

AAR20911  
ID AAR20911 standard; protein; 35 AA.

XX AAR20911;

XX 01-JUN-1992 (first entry)

XX PND EE533-2.

XX HIV; PND; principal neutralising determinant; Omp; Neisseria;  
XX outer membrane proteosome; AIDS; vaccine; envelope gene.

XX Human immunodeficiency virus.

XX EP471407-A.

XX 19-FEB-1992.

XX 07-AUG-1991; 91EP-00202025.

XX 13-AUG-1990; 90US-00566638.

XX 13-AUG-1990; 90US-00566654.

XX 13-AUG-1990; 90US-00566656.

XX (MERI ) MERCK & CO INC.

XX Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21335.

XX New antigenic conjugate of HIV major neutralisation determinant -

XX covalently linked to outer membrane proteosome of Neisseria, useful as

XX vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 140; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
XX encode PNDs of HIV (envelope protein) which are useful as immunogens for  
XX AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
XX conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,  
XX then linked by cross-linking agents, monogeneric spacers or bigeneric  
XX spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
XX or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
XX Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
XX post- exposure to prevent or treat HIV infection or disease, and are  
XX capable of producing an immune response specific for the immunogen. See  
XX also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 91.0%; Score 183; DB 2; Length 35;  
Best Local Similarity 91.4%; Pred. No. 4.2e-17;  
Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 35  
|:|||||||||:|||||||||:|||||||||:|||||  
Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAH 35

RESULT 9

AAR21135  
ID AAR21135 standard; protein; 35 AA.

XX AAR21135;

[illegible]

```

19-FEB-1992.
07-AUG-1991; 91EP-00202025.
13-AUG-1990; 90US-00566638.
13-AUG-1990; 90US-00566654.
13-AUG-1990; 90US-00566656.
(MERI ) MERCK & CO INC.
Lewis JA, Davide JP, Waterbury JA;
WPI; 1992-058471/08.
N-PSDB; AAQ21336.
New antigenic conjugate of HIV major proteolysation determinant -
covalently linked to outer membrane proteosome of Neisseria, useful as
vaccine and in treating and preventing HIV infections, AIDS and ARC.
Claim 1; Page 175 + 141; 177pp; English.
The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
encode PNDs of HIV (envelope protein) which are useful as immunogens for
AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
then linked by cross-linking agents, monogeneric spacers or bigeneric
spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
post- exposure to prevent or treat HIV infection or disease, and are
capable of producing an immune response specific for the immunogen. See
also AAQ21052-61, AAQ21194-311 and AAQ21321-78
Sequence 35 AA;
Query Match 91.0%; Score 183; DB 2; Length 35;
Best Local Similarity 91.4%; Pred. No. 4.2e-17;
Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CSRPNNTKRSIPMGGRAFYTTGGIIGDIROAHC 35
Db 1 CTRPNNTKRSIPMGGRAFYTTGGIIGDIROAHC 35
RESULT 11
AAR20910
ID AAR20910 standard; protein; 35 AA.
AC AAR20910;
XX
XX
DT 01-JUN-1992 (first entry)
XX
XX DE EE533-1.
XX
XX HIV; PND; principal neutralising determinant; Omp; Neisseria;
KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX
XX Human immunodeficiency virus.
OS
XX EP471407-A.
XX
XX 19-FEB-1992.
XX
XX 07-AUG-1991; 91EP-00202025.
XX
XX 13-AUG-1990; 90US-00566638.
XX 13-AUG-1990; 90US-00566654.
XX 13-AUG-1990; 90US-00566656.
XX (MERI ) MERCK & CO INC.
XX
XX Lewis JA, Davide JP, Waterbury JA;
XX

```

DR WPI; 1992-058471/08.  
 DR N-PSDB; AAQ21334.  
 XX  
 PT New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.  
 XX  
 PS Claim 1; Page 175 + 139-140; 177pp; English.  
 XX  
 CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78  
 XX  
 SQ Sequence 35 AA;

Query Match 91.0%; Score 183; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 4.2e-17;  
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGRAFVYTTGQIIIGDIRQAH 35  
 |:|||||||:|||||||:|||||||:|||||||:|||||||  
 Db 1 CTRPNNTKRSIPGPGRAFVYTTGDIIGDIRQAH 35

RESULT 12  
 AAR21136  
 ID AAR21136 standard; protein; 35 AA.  
 AC AAR21136;

XX 01-JUN-1992 (first entry)  
 DT PND BR228-3.  
 DE  
 XX HIV; PND; principal neutralising determinant; Omp; Neisseria;  
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.  
 KW Human immunodeficiency virus.

OS  
 XX EP471407-A.  
 XX  
 PD 19-FEB-1992.  
 XX  
 XX 07-AUG-1991; 91EP-00202025.  
 XX  
 PR 13-AUG-1990; 90US-00566638.  
 PR 13-AUG-1990; 90US-00566654.  
 PR 13-AUG-1990; 90US-00566656.  
 XX  
 XX (MERI ) MERCK & CO INC.  
 XX  
 XX Lewis JA, Davide JP, Waterbury JA;  
 PI  
 XX WPI; 1992-058471/08.  
 DR N-PSDB; AAQ21234.  
 DR

XX New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.  
 XX  
 PS Claim 1; Page 175 + 66; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic

CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78  
 XX  
 SQ Sequence 35 AA;

Query Match 91.0%; Score 183; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 4.2e-17;  
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFVYTTGQIIIGDIRQAH 35  
 |:|||||||:|||||||:|||||||:|||||||:|||||||  
 Db 1 CTRPNNTKRSIPGPGRAFVYTTGDIIGDIRQAH 35

RESULT 13  
 AAR29113  
 ID AAR29113 standard; protein; 35 AA.  
 XX  
 AC AAR29113;

XX 25-MAR-2003 (revised)  
 DT 02-APR-1993 (first entry)  
 XX  
 DE Group II HIV gp120 PND peptide 8926C.

XX Principal Neutralising determinant; Human immunodeficiency virus;  
 KW vaccine; Robson's analytical method; polymerase chain reaction;  
 KW Garnier-Osguthorpe-Robson method; GOR method; secondary structure.  
 XX Human immunodeficiency virus.

XX  
 FH Key Location/Qualifiers  
 FT Region 8..14  
 FT /note= "classification of PND peptides is determined by  
 FT the predicted secondary structure of this region - see  
 FT comments"  
 FT Region 15..18  
 FT /note= "conserved PND motif"

FT EP516135-A2.  
 XX  
 XX 02-DEC-1992.  
 XX  
 XX 29-MAY-1992; 92EP-00109072.  
 XX  
 XX 31-MAY-1991; 91JP-00129224.  
 XX  
 XX (KAGA ) CHERO SERO THERAPEUTIC RES INS.  
 XX  
 XX Eda Y, Shiosaki K, Osatomi K, Tokiyoshi S;  
 PI  
 XX WPI; 1992-400517/49.  
 DR  
 XX Principle neutralising determinant peptide(s) of HIV gp120 protein - used  
 PT for diagnosing, preventing and treating HIV infection.  
 PT  
 XX Example 1; Page 10; 26pp; English.

XX DNA encoding HIV PND peptides was PCR amplified using genomic DNA from  
 CC HIV-infected peripheral blood mononuclear cells as template. The  
 CC amplified fragments were fused to beta-galactosidase coding sequence.  
 CC E.coli transformants were cultured to produce the fusion protein. The  
 CC expressed PND proteins were divided into groups based on their reactivity  
 CC with neutralising antibodies and their amino acid sequence. The amino  
 CC acid sequence was analysed using Robson's analytical program for protein  
 CC secondary structure. Five groups were identified and 90% of all  
 CC previously reported PND peptides were included in 3 main groups (i.e.

Search completed: April 11, 2005, 11:09:02  
Job time : 14.7951 secs

**THIS PAGE LEFT BLANK**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:57:57 ; Search time 2.72785 Seconds  
(without alignments)  
1269.790 Million cell updates/sec

Title: US-09-762-261C-1\_COPY\_301\_336  
Perfect score: 201  
Sequence: 1 CSRPNNTKRSIPMGPGAFYTTGGIIGDIRQAHCN 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	90.5	77	2 S35862	env polyprotein -
2	182	90.5	110	2 A46410	env polyprotein -
3	182	90.5	847	2 T09448	env polyprotein -
4	182	90.5	847	2 S13289	env polyprotein -
5	181	90.0	76	2 S35859	env polyprotein -
6	181	90.0	77	2 S35820	env polyprotein -
7	181	90.0	77	2 S35819	env polyprotein -
8	180	89.6	35	2 PC2296	V3 domain peptide
9	180	89.6	443	2 C41621	env polyprotein P
10	179	89.1	77	2 S35834	env polyprotein -
11	179	89.1	77	2 S35836	env polyprotein -
12	179	89.1	77	2 S35835	env polyprotein -
13	178	88.6	77	2 S35848	env polyprotein -
14	178	88.6	77	2 S35849	env polyprotein -
15	178	88.6	77	2 S35857	env polyprotein -
16	177	88.1	46	2 G01847	env polyprotein -
17	177	88.1	77	2 S35837	env polyprotein -
18	177	88.1	77	2 S35809	env polyprotein -
19	177	88.1	110	2 B46410	env polyprotein -
20	177	88.1	110	2 C46410	env polyprotein -
21	177	88.1	290	2 S25940	env polyprotein -
22	177	88.1	445	2 A41621	env polyprotein M
23	176	87.6	35	2 PC2295	V3 domain peptide
24	176	87.6	77	2 S35806	env polyprotein -
25	176	87.6	77	2 S35841	env polyprotein -
26	176	87.6	77	2 S35816	env polyprotein -
27	176	87.6	77	2 S35804	env polyprotein -
28	176	87.6	77	2 S35850	env polyprotein -
29	176	87.6	843	1 H44001	env polyprotein pr

30	175	87.1	71	2 S42918	env polyprotein -
31	175	87.1	77	2 S35821	env polyprotein -
32	174	86.6	77	2 S35824	env polyprotein -
33	174	86.6	77	2 S35825	env polyprotein -
34	174	86.6	110	2 D46410	env polyprotein -
35	174	86.6	852	1 VCLJBR	env polyprotein -
36	173	86.1	71	2 S42904	env polyprotein -
37	173	86.1	71	2 S42905	env polyprotein -
38	173	86.1	71	2 S42898	env polyprotein -
39	173	86.1	77	2 S35815	env polyprotein -
40	172	85.6	506	2 A40218	env polyprotein -
41	171	85.1	35	2 PC2297	V3 domain peptide
42	171	85.1	77	2 S35800	env polyprotein -
43	171	85.1	77	2 S35801	env polyprotein -
44	171	85.1	299	2 S60529	env polyprotein -
45	171	85.1	855	1 VCLJAJ	env polyprotein pr

ALIGNMENTS

RESULT 1

S35862  
env polyprotein - human immunodeficiency virus type 1 (strain CSF2951) (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S35862; S35860; S35863

R;Chiodi, F.  
submitted to the EMBL Data Library, June 1993

A;Reference number: S35800

A;Accession: S35862

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-77 <CH1>

A;Cross-references: UNIPROT:Q77544; UNIPROT:Q77542; UNIPROT:Q77545; EMBL:Z23246; NID:g311

A;Experimental source: strain CSF2951 clone3; strain CSF2951 clone4

A;Accession: S35860

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-77 <CH2>

A;Cross-references: EMBL:Z23244; NID:g313631; PIDN:CAA80762.1; PID:g313632; EMBL:Z23247;

A;Experimental source: strain CSF2951 clone11; strain CSF2951 clone4

C;Superfamily: type E retrovirus env polyprotein

Query Match 90.5%; Score 182; DB 2; Length 77;

Best Local Similarity 88.9%; Pred. No. 2.8e-18;

Matches 32; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CSRPNNTKRSIPMGPGAFYTTGGIIGDIRQAHCN 36

Db 26 CTRPNNTKRSINIGPGAFYTTGGIIGDIRQAHCN 61

RESULT 2

A46410  
env polyprotein (V3 domain) - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A46410

R;McNairney, T.; Hornickova, Z.; Markham, R.; Birdwell, A.; Arens, M.; Saah, A.; Ratner,

Proc. Natl. Acad. Sci. U.S.A. 89, 10247-10251, 1992

A;Title: Relationship of human immunodeficiency virus type 1 sequence heterogeneity to st

A;Reference number: A46410; MUID:93066216; PMID:1438212

A;Accession: A46410

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-110 <MCN>

A;Cross-references: UNIPROT:Q9PY06

A;Experimental source: subject S1

A;Note: sequence extracted from NCBI backbone (NCBIP:117724)

C;Superfamily: type E retrovirus env polyprotein

Query Match 90.5%; Score 182; DB 2; Length 110;



Db 26 CTRPNNNTRKSIHMGPGRAFYATGDIIGDIRQAHCN 61

RESULT 11

S35836

envelope protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S35836

R:Chiodi, F.

submitted to the EMBL Data Library, June 1993

A:Reference number: S35800

A:Accession: S35836

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-77 <CHI>

A:Cross-references: UNIPROT:Q77518; EMBL:Z23218

C:Superfamily: type E retrovirus env polypeptide

Query Match 89.1%; Score 179; DB 2; Length 77;

Best Local Similarity 88.9%; Pred. No. 7.4e-18;

Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps

Qy 1 CSRPNNTKSIHMGPGRAFYATGDIIGDIRQAHCN 36

Db 26 CTRPNNNTRKSIHMGPGRAFYATGDIIGDIRQAHCN 61

RESULT 12

S35835

envelope protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S35835

R:Chiodi, F.

submitted to the EMBL Data Library, June 1993

A:Reference number: S35800

A:Accession: S35835

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-77 <CHI>

A:Cross-references: UNIPROT:Q77517; EMBL:Z23217

C:Superfamily: type E retrovirus env polypeptide

Query Match 89.1%; Score 179; DB 2; Length 77;

Best Local Similarity 88.9%; Pred. No. 7.4e-18;

Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps

Qy 1 CSRPNNTKSIHMGPGRAFYATGDIIGDIRQAHCN 36

Db 26 CTRPNNNTRKSIHMGPGRAFYATGDIIGDIRQAHCN 61

RESULT 13

S35848

envelope protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S35848

R:Chiodi, F.

submitted to the EMBL Data Library, June 1993

A:Reference number: S35800

A:Accession: S35848

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-77 <CHI>

A:Cross-references: UNIPROT:Q77530; EMBL:Z23232; NID:G313607; PIDN:CAA80750

C:Superfamily: type E retrovirus env polypeptide

Query Match 88.6%; Score 178; DB 2; Length 77;

Best Local Similarity 86.1%; Pred. No. 1e-17;

Matches 31; Conservative 3; Mismatches 2; Indels 0; Gaps

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36  
|:||||| |:||||| |:||||| |:|||||  
Db 26 CTRPNNTKRGHIGPGAFYTTGELIIGDIRQAHCN 61

## RESULT 14

S35849  
envelope protein - human immunodeficiency virus type 1 (fragment)  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004  
C;Accession: S35849  
R;Chiodi, F.  
submitted to the EMBL Data Library, June 1993  
A;Reference number: S35800  
A;Accession: S35849  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-77 <CHI>  
A;Cross-references: UNIPROT:Q77531; EMBL:Z23233; NID:g313609; PIDN:CAA80751.1; PID:g3136  
C;Superfamily: type E retrovirus env polyprotein

Query Match 88.6%; Score 178; DB 2; Length 77;  
Best Local Similarity 86.1%; Pred. No. 1e-17;  
Matches 31; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36  
|:||||| |:||||| |:||||| |:|||||  
Db 26 CTRPNNTRESIHIGPGAFYTTGELIIGDIRQAHCN 61

## RESULT 15

S35857  
envelope protein - human immunodeficiency virus type 1 (fragment)  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004  
C;Accession: S35857  
R;Chiodi, F.  
submitted to the EMBL Data Library, June 1993  
A;Reference number: S35800  
A;Accession: S35857  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-77 <CHI>  
A;Cross-references: UNIPROT:Q77539; EMBL:Z23241; NID:g313625; PIDN:CAA80759.1; PID:g3136  
C;Superfamily: type E retrovirus env polyprotein

Query Match 88.6%; Score 178; DB 2; Length 77;  
Best Local Similarity 86.1%; Pred. No. 1e-17;  
Matches 31; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36  
|:||||| |:||||| |:||||| |:|||||  
Db 26 CTRPNNTKRSWIGPGAFYTTGDIIGDIRQAHCN 61

Search completed: April 11, 2005, 11:15:28  
Job time : 2.72785 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:39:11 ; Search time 11.5261 Seconds  
(without alignments)  
1599.394 Million cell updates/sec

Title: US-09-762-261C-1\_COPY\_301\_336

Perfect score: 201  
Sequence: 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	866	2 Q9WPZ4	Q9WPZ4 human immun
2	198	98.5	91	2 Q39362	Q39362 human immun
3	198	98.5	91	2 Q69692	Q69692 human immun
4	195	97.0	85	2 Q90956	Q90956 human immun
5	195	97.0	88	2 Q9DXK1	Q9DXK1 human immun
6	195	97.0	171	2 Q6UA59	Q6UA59 human immun
7	194	96.5	92	2 Q90205	Q90205 human immun
8	194	96.5	92	2 Q90206	Q90206 human immun
9	194	96.5	92	2 Q79067	Q79067 human immun
10	194	96.5	112	2 Q9DPV9	Q9DPV9 human immun
11	194	96.5	112	2 Q9DPW0	Q9DPW0 human immun
12	194	96.5	112	2 Q9DPW1	Q9DPW1 human immun
13	194	96.5	112	2 Q9DPW2	Q9DPW2 human immun
14	194	96.5	112	2 Q9DPW3	Q9DPW3 human immun
15	194	96.5	113	2 Q71020	Q71020 human immun
16	194	96.5	229	2 Q40047	Q40047 human immun
17	194	96.5	229	2 Q40049	Q40049 human immun
18	194	96.5	229	2 Q40051	Q40051 human immun
19	194	96.5	230	2 Q40055	Q40055 human immun
20	194	96.5	235	2 Q40050	Q40050 human immun
21	194	96.5	235	2 Q40052	Q40052 human immun
22	193	96.0	80	2 Q9QG77	Q9QG77 human immun
23	193	96.0	119	2 Q75417	Q75417 human immun
24	193	96.0	119	2 Q75422	Q75422 human immun
25	193	96.0	119	2 Q75426	Q75426 human immun
26	193	96.0	120	2 Q75415	Q75415 human immun
27	193	96.0	120	2 Q75421	Q75421 human immun
28	193	96.0	170	2 Q8AEX6	Q8AEX6 human immun
29	193	96.0	311	2 Q72513	Q72513 human immun
30	192	95.5	35	2 Q78198	Q78198 human immun
31	192	95.5	74	2 Q8A1V5	Q8A1V5 human immun

RESULT 1  
Q9WPZ4 PRELIMINARY; PRT; 866 AA.  
AC Q9WPZ4;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Envelope protein.  
OS Name=env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retrovirus; Retroviridae; Lentivirus.  
OC NCBI\_TaxID=11676;  
RX [1]  
RX MEDLINE=99236722; PubMed=10221533; DOI=10.1089/088922299311088;  
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;  
RT "Expression and characterization of HIV type 1 envelope protein associated with a broadly reactive neutralizing antibody response.";  
RL AIDS Res. Hum. Retroviruses 15:561-570(1999).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=99236722; PubMed=10221533; DOI=10.1089/088922299311088;  
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF128126; AD40637.3; -.  
DR HSSP; P04578; 1DLB.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 866 AA; 98113 MW; 6E76021833P2EACD CRC64;

Query Match 100.0%; Score 201; DB 2; Length 866;  
Best Local Similarity 100.0%; Pred. No. 1e-19;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 36  
|||||  
Db 301 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 336

RESULT 2  
Q39362 PRELIMINARY; PRT; 91 AA.  
ID Q39362;  
AC Q39362;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).

32 192 95.5 74 2 Q8A1W7  
33 192 95.5 111 2 Q9Q8E5  
34 192 95.5 198 2 Q6SY25  
35 192 95.5 198 2 Q6SY26  
36 192 95.5 198 2 Q6SY27  
37 192 95.5 198 2 Q6SY28  
38 192 95.5 198 2 Q6SY29  
39 191 95.0 53 2 Q9PXN8  
40 191 95.0 54 2 Q76143  
41 191 95.0 68 2 Q9QJ11  
42 191 95.0 73 2 Q92991  
43 191 95.0 73 2 Q92992  
44 191 95.0 75 2 Q90U52  
45 191 95.0 75 2 Q90U56  
Q8A1W7 human immun  
Q9Q8E5 human immun  
Q6SY25 human immun  
Q6SY26 human immun  
Q6SY27 human immun  
Q6SY28 human immun  
Q6SY29 human immun  
Q9PXN8 human immun  
Q76143 human immun  
Q9QJ11 human immun  
Q92991 human immun  
Q92992 human immun  
Q90U52 human immun  
Q90U56 human immun

#### ALIGNMENTS

```

GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98242906; PubMed=983604;
RA Lukashov V.V.; Op de Coul E.L.; Coutinho R.A.; Goudsmit J.;
RT "HIV-1 strains specific for Dutch injecting drug users in
RT heterosexually infected individuals in The Netherlands.";
RL AIDS 12:635-641(1998).
DR EMBL; AF032157; AAC59382.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 91
FT NON_TER 91 91
SQ SEQUENCE 91 AA; 10363 MW; 0D9690FA2767592E CRC64;

Query Match 98.5%; Score 198; DB 2; Length 91;
Best Local Similarity 97.2%; Pred. No. 2.4e-20;
Matches 35; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIKPMGPGRAFYTGTGIIIGDIRQAHCN 36
Db [1]
27 CTRPNNTKSIKPMGPGRAFYTGTGIIIGDIRQAHCN 62

RESULT 3
Q69692 PRELIMINARY; PRT; 91 AA.
AC Q69692;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=R923572;
RA Lukashov V.V.; Kuiken C.L.; Boer K.; Goudsmit J.;
RT "HIV type 1 subtypes in The Netherlands circulating among women
RT originating from AIDS-endemic regions.";
RL AIDS Res. Hum. Retroviruses 12:951-953(1996).
DR EMBL; L76886; AAC37946.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 91
FT NON_TER 91 91
SQ SEQUENCE 91 AA; 10333 MW; 17273BE096D7593E CRC64;

Query Match 98.5%; Score 198; DB 2; Length 91;
Best Local Similarity 97.2%; Pred. No. 2.4e-20;
Matches 35; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIKPMGPGRAFYTGTGIIIGDIRQAHCN 36
Db [1]
27 CTRPNNTKSIKPMGPGRAFYTGTGIIIGDIRQAHCN 62

```

```

RESULT 4
O90956 PRELIMINARY; PRT; 85 AA.
AC O90956;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp120 (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Heyndrickx L.; Janssens W.; Coppens S.; Vereecken K.; Willems B.;
RA Franssen K.; Colebunders R.; Vandendruene M.; Van der Groen G.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ228189; CAAL2819.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Transmembrane.
FT NON_TER 1 85
FT NON_TER 85 85
SQ SEQUENCE 85 AA; 9889 MW; 01BB36F634EA0AFB CRC64;

Query Match 97.0%; Score 195; DB 2; Length 85;
Best Local Similarity 94.4%; Pred. No. 5.9e-20;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIKPMGPGRAFYTGTGIIIGDIRQAHCN 36
Db [1]
29 CTRPNNTKSIKPMGPGRAFYTGTGIIIGDIRQAHCN 64

RESULT 5
Q9DXK1 PRELIMINARY; PRT; 88 AA.
AC Q9DXK1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21084776; PubMed=11216936;
RA Op de Coul E.L.M.; Prins M.; Cornelissen M.; van der Schoot A.;
RA Boufassa F.; Brettelle R.P.; Hernandez-Aguado L.; Schiffer V.;
RA McMenamin J.; Rezza G.; Robertson R.; Zangerle R.; Goudsmit J.;
RA Coutinho R.A.; Lukashov V.V.;
RT Using phylogenetic analysis to trace HIV-1 migration among western
RT European injecting drug users seroconverting from 1984 to 1997.";
RL AIDS 15:257-266(2001).
DR EMBL; AF307295; AAG32463.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 88
FT NON_TER 88 88
SQ SEQUENCE 88 AA; 9925 MW; D523F6D428DB9E4 CRC64;

Query Match 97.0%; Score 195; DB 2; Length 88;

```

KW	AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT	NON TER 1
FT	NON TER 92
SQ	SEQUENCE 92 AA; 10466 MW; 3395C4528B9A2E51 CRC64;
	Query Match 96.5%; Score 194; DB 2; Length 92;
	Best Local Similarity 94.4%; Pred. No. 9e-20;
	Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY	1 CSRPNNTKRSIPMGPGRAFYYTGGIIGDIRQAHCN 36
	:-:     :-:     :-:     :-:     :-:
Db	27 CTRPNNTKRSIPMGPGRAFYYTGGIIGDIRQAHCN 62
RESULT 8	
ID	O90206 PRELIMINARY; PRT; 92 AA.
AC	O90206;
DT	01-NOV-1998 (TrEMBLrel. 08, Created)
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Envelope glycoprotein (Fragment).
GN	Name=env;
OS	Human immunodeficiency virus 1.
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX	NCBI_TaxID=11676;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99039932; PubMed=9824326;
RA	Goudmit J., Lukashov V.V., van Ameijden E.J., Zorgdrager F.,
RA	van den Burg R., Cornelissen M.;
RT	"Impact of sexual versus parenteral transmission events on the
RT	evolution of the gag and env genes of HIV type 1.";
RL	AIDS Res. Hum. Retroviruses 14:1483-1486(1998).
DR	EMBL; AF071286; AAC71738.1; -.
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0019028; C:viral capsid; IEA.
DR	GO; GO:0019031; C:viral envelope; IEA.
DR	GO; GO:0005198; F:structural molecule activity; IEA.
DR	InterPro; IPR000777; GP120.
DR	InterPro; IPR007110; Ig-like.
DR	Pfam; PF00516; GP120; 1.
KW	AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT	NON TER 1
FT	NON TER 92
SQ	SEQUENCE 92 AA; 10482 MW; 3395DB8810C080F1 CRC64;
	Query Match 96.5%; Score 194; DB 2; Length 92;
	Best Local Similarity 94.4%; Pred. No. 9e-20;
	Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY	1 CSRPNNTKRSIPMGPGRAFYYTGGIIGDIRQAHCN 36
	:-:     :-:     :-:     :-:     :-:
Db	27 CTRPNNTKRSIPMGPGRAFYYTGGIIGDIRQAHCN 62
RESULT 9	
ID	Q79067 PRELIMINARY; PRT; 92 AA.
AC	Q79067;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Envelope protein (Fragment).
GN	Name=env;
OS	Human immunodeficiency virus 1.
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX	NCBI_TaxID=11676;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=94022315; PubMed=8415653;
RA	Kuiken C.L., Zwart G., Baan E., Coutinho R.A., Hoek van den J.A.R.,
RA	Goudmit J.;

RT "Increasing antigenic and genetic diversity of the HIV-1 V3 domain in  
 the course of the AIDS epidemic";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:9065-9065(1993).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=96203966; PubMed=8627267;  
 RA Kuiken C.L., Cornelissen M.T.E., Zorgrader F., Hartman S.,  
 RA Gibbs A.J., Goudsmit J.;  
 RT "Consistent risk group-associated differences in human  
 RT immunodeficiency virus type 1 vpr, vpu and V3 sequences despite  
 RT independent evolution";  
 RL J. Gen. Virol. 77:783-792(1996).  
 RN [3]

RP SEQUENCE FROM N.A.  
 RA Kuiken L.;  
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z29323; CA82520.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000777; GP120.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 92 92  
 SQ SEQUENCE 92 AA; 10443 MW; D227D77BDB2283E5 CRC64;

Query Match 96.5%; Score 194; DB 2; Length 92;  
 Best Local Similarity 94.4%; Pred. No. 9e-20;  
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 26 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 61  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 10

ID Q9DPV9 PRELIMINARY; PRT; 112 AA.  
 AC Q9DPV9;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wang Y.M., Saksena N.K.;  
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1  
 RT quasiespecies in a patient with periodic changes to the HAART  
 RT regimens";  
 RL AIDS 0:0-0(2001).  
 DR EMBL; AF291142; AAG35348.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 112 112  
 SQ SEQUENCE 112 AA; 12603 MW; A2F0458A21E5C6E2 CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;  
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;  
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 11

ID Q9DPW0 PRELIMINARY; PRT; 112 AA.  
 AC Q9DPW0;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wang Y.M., Saksena N.K.;  
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1  
 RT quasiespecies in a patient with periodic changes to the HAART  
 RT regimens";  
 RL AIDS 0:0-0(2001).  
 DR EMBL; AF291141; AAG35347.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 112 112  
 SQ SEQUENCE 112 AA; 12587 MW; 16B8F2D91CBB6AEE CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;  
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;  
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 12

ID Q9DPW1 PRELIMINARY; PRT; 112 AA.  
 AC Q9DPW1;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wang Y.M., Saksena N.K.;  
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1  
 RT quasiespecies in a patient with periodic changes to the HAART  
 RT regimens";  
 RL AIDS 0:0-0(2001).  
 DR EMBL; AF291140; AAG35346.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 112 112  
 SQ SEQUENCE 112 AA; 12587 MW; 16B8F2D91CBB6AEE CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;  
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;  
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36  
 Db 28 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 63  
 [2] :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||



```

DR GO: 0019028; C:viral capsid; IEA.
DR GO: 0019031; C:viral envelope; IEA.
DR GO: 0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR AIDS: Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 112
FT NON_TER 113
SQ SEQUENCE 112 AA; 12586 MW; A2E8F2D91CBB7034 CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;
Best Local Similarity 94.4%; Pred. No. 1.1e-19;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYYTTGQIIIGDIRQAHCN 36
   |:|||||:|||||:|||||:|||||:|||||
Db 28 CTRPNNTKRSIPGPGRAFYYTTGQIIIGDIRQAHCN 63
   |:|||||:|||||:|||||:|||||:|||||

RESULT 15
Q71020 PRELIMINARY; PRT; 113 AA.
AC Q71020;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein, C2-V3 region (Fragment).
DE Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95194719; PubMed=7888212;
RA Kalish M.L., Luo C.C., Weniger B.G., Limpakarnjanarat K., Young N.,
RA Ou C.Y., Schochetman G.;
RA "Early HIV type 1 strains in Thailand were not responsible for the
RT current epidemic.";
RL AIDS Res. Hum. Retroviruses 10:1573-1575(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Yesley A.R.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U15579; AAC54014.1; -.
DR GO: 0016021; C:integral to membrane; IEA.
DR GO: 0019028; C:viral capsid; IEA.
DR GO: 0019031; C:viral envelope; IEA.
DR GO: 0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR AIDS: Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 113
FT NON_TER 113
SQ SEQUENCE 113 AA; 12642 MW; B1DBE4BC87669819 CRC64;

Query Match 96.5%; Score 194; DB 2; Length 113;
Best Local Similarity 94.4%; Pred. No. 1.1e-19;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYYTTGQIIIGDIRQAHCN 36
   |:|||||:|||||:|||||:|||||:|||||
Db 19 CTRPNNTKRSIPGPGRAFYYTTGQIIIGDIRQAHCN 54
   |:|||||:|||||:|||||:|||||:|||||

Search completed: April 11, 2005, 11:14:04
Job time : 12.5261 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:37:31 ; Search time 283.737 Seconds  
(without alignments)  
1180.438 Million cell updates/sec

Title: US-09-762-261C-1  
Perfect score: 4517  
Sequence: 1 MRVKGIRNRYQHNGWGTML.....GRALLHIPTRIOGLERALL 866

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4596	99.5	866	3 AAY79020	Aay79020 HIV-1 env
2	3971.5	86.0	857	2 AAR14904	Aar14904 HIV-1 (MN-
3	3971.5	86.0	857	2 AAW11580	Aaw11580 Human Imm
4	3971.5	86.0	857	2 AAW89112	Aaw89112 Predicted
5	3970	86.0	856	2 AAW11579	Aaw11579 Human Imm
6	3970	86.0	856	2 AAW88111	Aaw88111 Predicted
7	3941.5	85.4	857	8 ADP04371	Adp04371 Envelope
8	3935	85.2	856	2 AAW58805	Aaw58805 HIV-1 iso
9	3933	85.2	856	2 AAR79173	Aar79173 HIV virus
10	3920.5	84.9	883	4 AAB82761	Aab82761 Ancestral
11	3917	84.8	856	2 AAR14903	Aar14903 HIV-1 (MN)
12	3913	84.8	854	2 AAW43068	Aaw43068 HIV-1 gpl
13	3896	84.4	856	6 ABR55495	Abr55495 Amino aci
14	3859.5	83.6	863	8 ADP04369	Adp04369 Envelope
15	3845.5	83.3	847	8 ADK14406	Adk14406 HIV wild-
16	3845.5	83.3	847	8 ADP20081	Adp20081 Human imm
17	3841.5	83.2	847	8 ADK14404	Adk14404 HIV mutan
18	3841	83.2	850	2 AAR67724	Aar67724 gp120 fro
19	3839.5	83.2	847	8 ADK14405	Adk14405 HIV mutan
20	3814	82.6	848	8 ADP20072	Adp20072 Human imm
21	3812.5	82.6	855	2 AAW11581	Aaw11581 Human imm
22	3812.5	82.6	855	2 AAW88113	Aaw88113 Env prote
23	3806.5	82.4	857	2 AAR67725	Aar67725 gp120 fro
24	3795.5	82.2	847	3 AAY97073	Aay97073 Variant H
25	3795.5	82.2	847	8 ADO05103	Ado05103 Human imm

26	3787.5	82.0	851	1 AAP80967	Aap80967 HIV prote
27	3768	81.6	856	2 AAR67726	Aar67726 gp120 fro
28	3758.5	81.4	855	2 AAW43069	Aaw43069 HIV-1 gpl
29	3753.5	81.3	868	5 AAO19389	Aao19389 Lymphaden
30	3744	81.1	856	2 AAR25940	Aar25940 Modified
31	3744	81.1	856	4 AAB85999	Aab85999 Amino aci
32	3744	81.1	863	2 AAR43869	Aar43869 HTLV-III
33	3738.5	81.0	855	8 ADO52541	Ado52541 HIV-1 rec
34	3735.5	80.9	868	1 AAP60422	Aap60422 Sequence
35	3734.5	80.9	868	1 AAP60063	Aap60063 HIV virus
36	3734	80.9	863	1 AAP60349	Aap60349 HTLV-III
37	3733.5	80.9	901	1 AAP70665	Aap70665 Sequence
38	3732.5	80.8	849	8 ADO52531	Ado52531 HIV-1 rec
39	3731	80.8	856	2 AAW89325	Aaw89325 HIV-1 env
40	3731	80.8	856	6 ABU63322	Abu63322 Human lym
41	3731	80.8	856	8 ADO26434	Ado26434 HTLV-III
42	3726.5	80.7	848	5 AAU11872	Aau11872 HIV env p
43	3726	80.7	856	3 AAY97072	Aay97072 Wild type
44	3726	80.7	856	8 ADE84721	Ade84721 Human imm
45	3726	80.7	856	8 ADN36425	Adn36425 HIV prote

## ALIGNMENTS

RESULT 1  
AAY79020  
ID AAY79020 standard; protein; 866 AA.  
XX AC AAY79020;  
XX DT 12-SEP-2003 (revised)  
DT 05-JUN-2000 (first entry)  
XX DE HIV-1 envelope protein amino acid sequence.  
XX KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;  
KW vaccine; antibody; detect.  
XX OS Human immunodeficiency virus 1.  
XX PN W0200007631-A1.  
XX PD 17-FEB-2000.  
XX PF 04-AUG-1999; 99WO-US017596.  
XX PR 04-AUG-1998; 98US-0095267P.  
XX PA (JACK-) JACKSON FOUND HENRY M.  
XX PI Quinnan GV, Zhang PF;  
XX DR WPI; 2000-2055578/18.  
XX PT Human immunodeficiency virus-1 envelope protein, useful as a vaccine or  
PT immunogenic preparation against HIV-1 infection induces the production of  
PT broadly cross-reactive neutralizing anti-serum.  
XX PS Claim 2; Page 33; 54pp; English.  
XX CC This sequence represents a human immunodeficiency virus type 1 (HIV-1)  
CC envelope (env) protein amino acid sequence. The invention relates to the  
CC HIV-1 env protein or its fragments, which when administered to a mammal,  
CC induces the production of broadly cross-reactive neutralising anti-serum  
CC against multiple strains of HIV-1. The HIV-1 env protein and its  
CC fragments are useful as vaccines, immunogenic compositions or diagnostic  
CC reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing  
CC antibodies in a sample. The amino acid sequence and its fragments or mRNA  
CC sequence are useful for generating antibodies in a mammal. In addition, a  
CC recombinant delivery vector containing the env amino acid sequence may  
CC also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS  
CC field)

```
XX SQ Sequence 866 AA;
Query Match 99.5%; Score 4596; DB 3; Length 866;
Best Local Similarity 99.5%; Pred. No. 2.6e-243;
Matches 862; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWTVVYGVVPWKEATTLFCASDAKAY 60
DB 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWTVVYGVVPWKEATTLFCASDAKAY 60

QY 61 DTEAHNVWATHACVPTDPPNPQEVVELNVVTENFNMWKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPPNPQEVVELNVVTENFNMWKNMVEQMHEDIISLWDSLKPCVK 120

QY 121 LTPCLVTLNCTDLRNTNTNNSTNNNSSEGTIKGEMKNCNFNIATSGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNSTNNNSSEGTIKGEMKNCNFNIATSGDKMQKEYAL 180

QY 181 LYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
DB 181 LYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240

QY 241 KGSKKNVSTVQCTHGRIPVSVSTQLLNGSLAEEVVRSENFTNNAKTIIVQLREPVKIN 300
DB 241 KGSKKNVSTVQCTHGRIPVSVSTQLLNGSLAEEVVRSENFTNNAKTIIVQLREPVKIN 300

QY 301 CSRPNNNTRKSIIPMGPGRAFYTTGQIIGDIRQAHCHNISKTNTNALKQVVEKLGEOFNKT 360
DB 301 CSRPNNNTRKSIIPMGPGRAFYTTGQIIGDIRQAHCHNISKTNTNALKQVVEKLGEOFNKT 360

QY 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNITQOLFDSIWNSENGTWNITRGLNNTGRNDTI 420
DB 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNITQOLFDSIWNSENGTWNITRGLNNTGRNDTI 420

QY 421 TLPCKRIQIINRQWGVKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
DB 421 TLPCKRIQIINRQWGVKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480

QY 481 GGDMDNRSELKYKVVKIEPLGVAPTAKRVVQREERAVGIGFPGAAGSTMG 540
DB 481 GGDMDNRSELKYKVVKIEPLGVAPTAKRVVQREERAVGIGFPGAAGSTMG 540

QY 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQITVWGIKQLOARILAVERYLKDQ 600
DB 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQITVWGIKQLOARILAVERYLKDQ 600

QY 601 LLGTWCGSGKLICTTVPWNASWSKNTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660
DB 601 LLGTWCGSGKLICTTVPWNASWSKNTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660

QY 661 IQOEKNEQELLELDKWANLWNPFDISNWLWYIKIFIMVGLVGLRIVFVVLISVNRVQ 720
DB 661 IQOEKNEQELLELDKWANLWNPFDISNWLWYIKIFIMVGLVGLRIVFVVLISVNRVQ 720

QY 721 GYSPFLSFQTRLPAPRGDPREEIEEGGDRDRSGLLVDGFTLIWDLRSCLFSYHR 780
DB 721 GYSPFLSFQTRLPAPRGDPREEIEEGGDRDRSGLLVDGFTLIWDLRSCLFSYHR 780

QY 781 LRDLILLIVTRIVELLGRGWEILLKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840
DB 781 LRDLILLIVTRIVELLGRGWEILLKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840

QY 841 QVLQRVGRALLHIPTRIROGLERALL 866
DB 841 QVLQRVGRALLHIPTRIROGLERALL 866

RESULT 2
AAR14904
ID AAR14904 standard; protein; 857 AA.
XX
AC AAR14904;
```

```
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-DEC-2001 (revised)
DT 05-FEB-1992 (first entry)
XX HIV-1 (MN-ST1) env protein.
DE human immunodeficiency virus; United States; MN isolate; AIDS;
KW envelope protein.
XX Human immunodeficiency virus 1.
OS USN7599491-N.
PN 15-OCT-1991.
PD 17-OCT-1990; 90US-00183830.
PF 17-OCT-1990; 90US-00599491.
PR (USSH ) NAT INST OF HEALTH.
PA Reitz M;
XX WPI; 1991-346752/47.
DR P-PSDB; AAR14904.
XX US HIV-1 isolates MN-ST1 and BA-L, ENV protein and DNA - are useful in
PT therapeutics, vaccines and diagnostic tests.
XX Example 2; Fig 6; 61pp; English.
CC The infectious molecular clone, lambda MN-ST1, was obtained by cloning
CC integrated provirus from DNA purified from peripheral blood lymphocytes
CC infected with HIV-1 (MN) and maintained in culture for one month. The
CC integrated proviral DNA was partially digested with Sau3A to give
CC fragments of 15-20 kb. The fragments were cloned in EMBL3 and the entire
CC sequence of the clone was determined and the amino acid sequence of the
CC env protein was deduced from it. (Note: Revised entry submitted to
CC correct the patent number format of US Government-owned NTIS applications
CC to prevent clashes with ongoing US granted patent numbers. For further
CC information please visit the Derwent web site at
CC www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 25-MAR-2003 to
CC correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 857 AA;
SQ
Query Match 86.0%; Score 3971.5; DB 2; Length 857;
Best Local Similarity 86.6%; Pred. No. 3.7e-209;
Matches 750; Conservative 37; Mismatches 70; Indels 9; Gaps 5;

QY 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWTVVYGVVPWKEATTLFCASDAKAY 60
DB 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWTVVYGVVPWKEATTLFCASDAKAY 60

QY 61 DTEAHNVWATHACVPTDPPNPQEVVELNVVTENFNMWKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPPNPQEVVELNVVTENFNMWKNMVEQMHEDIISLWDSLKPCVK 120

QY 121 LTPCLVTLNCTDLRNTNTNNSTNNNSSEGTIKGEMKNCNFNIATSGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNSTNNNSSEGTIKGEMKNCNFNIATSGDKMQKEYAL 180

QY 181 LYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
DB 181 LYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240

QY 241 KGSKKNVSTVQCTHGRIPVSVSTQLLNGSLAEEVVRSENFTNNAKTIIVQLREPVKIN 300
DB 241 KGSKKNVSTVQCTHGRIPVSVSTQLLNGSLAEEVVRSENFTNNAKTIIVQLREPVKIN 300

QY 301 CSRPNNNTRKSIIPMGPGRAFYTTGQIIGDIRQAHCHNISKTNTNALKQVVEKLGEOFNKT 360
```

```

Db      301 CTRPNYKRIHIGPGRFYTTKNIIGTIRQAHCHNISRAKWNLTLEQIVSKLKEQFKNK 360
Qy      361 KIVFTNSSGGDPEIVTHSFNCAGFEFFYCNNTQOLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db      361 TIVFNQSSGGDPEIVMHSFNCGGEFFYCNTPLFNSTWNG-NNTWN-----NTTGSNNNI 414
Qy      421 TLPCKRIKOIINRWQEVCKAMYPPIKGNISCSNITGLLITRDGKDDNSRDGNETFRPG 480
Db      415 TLQCKIKOIINMWQEVGKAIYAPPIEGQIRCSNITGLLITRDGKDDTND-TEIFRPG 473
Qy      481 GGDNRDNRSELYKYKVKIEPLGVAPTAKARRVVOREERAVGLGAMFFGLGAAGSTMG 540
Db      474 GGDNRDNRSELYKYKVTIEPLGVAPTAKARRVVOREKRA-AIGALFLGLGAAGSTMG 532
Qy      541 AASVTLTVOARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKOLQARILAVERYLKDDQ 600
Db      533 AASVTLTVOARQLLSGIVQOQSNLLRAIEAQHMLQLTVWGIKOLQARILAVERYLKDDQ 592
Qy      601 LLGIWCGSKGLICTTTVPWNASWSKNTLEAIWNMTWMQWDKEIDNYTSLIYSLIESP 660
Db      593 LLGIWCGSKGLICTTTVPWNASWS-NKSLDDIWNMTWMQWEREIDNYTSLIYSLLEKSQ 651
Qy      661 IQQEKNEQELLELDKWANLWNPDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720
Db      652 TQQEMNEQELLELDKWASLWNPFDITNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 711
Qy      721 GYSPLSQTRLPAPRGDRPEIEEGGDRDRDRSGLLVDGFLTLIWDVLRSLCLFSYHR 780
Db      712 GYSPLSLQTRPPVPRGDRPEIEEGGDRDRDRTSGRLVHGLAIIWVDLRSLFLSYHH 771
Qy      781 LRDLIIIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATATAVAEGTDRVI 840
Db      772 LRDLIIIAARIVELLGRRGWEILKYWNLLQYWSQELKSSAVSLNATDIAVAEGTDRVI 831
Qy      841 QVLQRGALLHIPTRIQGLERALL 866
Db      832 EVLQRAIRAILHIPTRIQGLERALL 857

RESULT 3
AAW11580
ID      AAW11580 standard; protein; 857 AA.
XX
AC      AAW11580;
XX
DT      17-OCT-2003 (revised)
DT      25-MAR-2003 (revised)
DT      25-MAR-1997 (first entry)
XX
XX      Human Immunodeficiency Virus-1 strain MN-ST1 envelope protein.
XX
XX      Acquired immune deficiency syndrome; AIDS; envelope protein; env gene;
XX      vaccine.
XX
XX      Human immunodeficiency virus 1; (strain MN-ST1).
XX
XX      US5576000-A.
XX
XX      19-NOV-1996.
XX
XX      15-FEB-1995; 95US-00388809.
XX
XX      17-OCT-1990; 90US-00599491.
XX
XX      25-FEB-1993; 93US-00022835.
XX
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX      Gallo RC, Franchini G, Lori FC, Popovic M, Reitz MS, Gartner S;
XX      Markham PD;
XX
XX      WFI; 1997-011206/01.
XX
XX      N-PSDB; AAT58550.

```

```

XX
PT      New isolated envelope protein of HIV-1 strain BA-L and recombinant
PT      equivalents - useful as immunogens for vaccines and antibody prodn.,
XX      typical of US clinical isolates.
XX
PS      Example; Fig 6; 86pp; English.
XX
CC      The infectious molecular clone, lambda MN-ST1, was obtained by cloning
CC      integrated provirus from DNA purified from peripheral blood lymphocytes
CC      infected with HIV-1 (MN) and maintained in culture for one month. The
CC      integrated proviral DNA was partially digested with Sau3A to produce
CC      fragments of 15-20 kb. The fragments were cloned into the BamHI site of
CC      lambda EMBL3. The DNA sequence of the entire clone was determined and the
CC      env gene was excised and cloned into an expression plasmid for
CC      recombinant production of the envelope protein. The env protein has the
CC      present sequence and can be used as an immunogen for raising antibodies
CC      against HIV. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 17
XX      -OCT-2003 to standardise OS field)
SQ      Sequence 857 AA;

```

```

Query Match      86.0%; Score 3971.5; DB 2; Length 857;
Best Local Similarity 86.6%; Pred. No. 3.7e-209;
Matches 750; Conservative 37; Mismatches 70; Indels 9; Gaps 5;

Qy      1 MRVKGIRRNQHWGWTMLLGLLMTCSAATEKLWTVVYGVVPWKEATTTILFCASDAKAY 60
Db      1 MRVKGIRRNQHWGWTMLLGLLMTCSAATEKLWTVVYGVVPWKEATTTILFCASDAKAY 60
Qy      61 DTEAHNVWATHACVPTDPNPQVEVLNVNTEFNWKNMVEQMHEDIISLWDSQLKPCVK 120
Db      61 DTEAHNVWATHACVPTDPNPQVEVLNVNTEFNWKNMVEQMHEDIISLWDSQLKPCVK 120
Qy      121 LTPCLVTNLCTDLRNTTNTNNSDNNNSSEGTIKGEMKNCSEFNATISGDKMKQKEYAL 180
Db      121 LTPCLVTNLCTDLRNTTNTNNSDNNNSSEGTIKGEMKNCSEFNATISGDKMKQKEYAL 180
Qy      181 LYKLDIEPTDNDNTSYRLISCNSTSVITQACPKISFEPPIPIHYCAPAGFAILKCNCKKFSG 240
Db      181 LYKLDIVSINNDSTSYRLISCNSTSVITQACPKISFEPPIPIHYCAPAGFAILKCNCKKFSG 240
Qy      241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVWIRSENFNTNNAKTIIVQLREPKIN 300
Db      241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVWIRSENFNTNNAKTIIVHLNESVQIN 300
Qy      301 CSRPNNTKRSIPMGPGRAFYTGTQIIGDIRQAHCHNISKTNTNNAKQVVEKLGGEQFNKT 360
Db      301 CTRPNYKRIHIGPGRFYTTKNIIGTIRQAHCHNISRAKWNLTLEQIVSKLKEQFKNK 360
Qy      361 KIVFTNSSGGDPEIVTHSFNCAGFEFFYCNNTQOLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db      361 TIVFNQSSGGDPEIVMHSFNCGGEFFYCNTPLFNSTWNG-NNTWN-----NTTGSNNNI 414
Qy      421 TLPCKRIKOIINRWQEVCKAMYPPIKGNISCSNITGLLITRDGKDDNSRDGNETFRPG 480
Db      415 TLQCKIKOIINMWQEVGKAIYAPPIEGQIRCSNITGLLITRDGKDDTND-TEIFRPG 473
Qy      481 GGDNRDNRSELYKYKVKIEPLGVAPTAKARRVVOREERAVGLGAMFFGLGAAGSTMG 540
Db      474 GGDNRDNRSELYKYKVTIEPLGVAPTAKARRVVOREKRA-AIGALFLGLGAAGSTMG 532
Qy      541 AASVTLTVOARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKOLQARILAVERYLKDDQ 600
Db      533 AASVTLTVOARQLLSGIVQOQSNLLRAIEAQHMLQLTVWGIKOLQARILAVERYLKDDQ 592
Qy      601 LLGIWCGSKGLICTTTVPWNASWSKNTLEAIWNMTWMQWDKEIDNYTSLIYSLIESP 660
Db      593 LLGIWCGSKGLICTTTVPWNASWS-NKSLDDIWNMTWMQWEREIDNYTSLIYSLLEKSQ 651
Qy      661 IQQEKNEQELLELDKWANLWNPDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720
Db      652 TQQEMNEQELLELDKWASLWNPFDITNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 711

```



XX Acquired immune deficiency syndrome; AIDS; envelope protein; env gene;  
 KW vaccine.  
 XX Human immunodeficiency virus 1; (strain MN-PH1).  
 OS  
 XX US5576000-A.  
 PN  
 XX 19-NOV-1996.  
 PD  
 XX 15-FEB-1995; 95US-00388809.  
 XX  
 XX 17-OCT-1990; 90US-00599491.  
 PR  
 XX 25-FEB-1993; 93US-00022835.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Gallo RC, Franchini G, Lori FC, Popovic M, Reitz MS, Gartner S;  
 PI Markham PD;  
 PI  
 XX WPI; 1997-011206/01.  
 DR N-PSDB; AAT58549.  
 XX  
 XX New isolated envelope protein of HIV-1 strain BA-L and recombinant  
 PT equivalents - useful as immunogens for vaccines and antibody prodn.,  
 PT typical of US clinical isolates.  
 XX  
 XX Example; Fig 3; 86pp; English.  
 PS  
 XX The permuted circular unintegrated viral DNA representing the complete  
 CC HIV-1(MN) genome was cloned into the EcoRI site of lambda gtWES.lambda B  
 CC DNA from total DNA of H9 cells producing HIV-1(MN) isolate. The clone was  
 CC designated lambda MN-PH1. The present sequence was deduced from the env  
 CC gene region of the MN-PH1 genome. (Updated on 25-MAR-2003 to correct PF  
 CC field.) (Updated on 17-OCT-2003 to standardise OS field)  
 CC  
 XX Sequence 856 AA;  
 SQ

Query Match 86.0%; Score 3970; DB 2; Length 856;  
 Best Local Similarity 86.7%; Pred. No. 4.5e-209;  
 Matches 751; Conservative 36; Mismatches 69; Indels 10; Gaps 6;  
 1 MRVKGIRRYQHWWGHTMLGLLMTCATEKLWVTYGVVPVWKEATTLFCASDAKAY 60  
 1 MRVKGIRRYQHWWGHTMLGLLMTCATEKLWVTYGVVPVWKEATTLFCASDAKAY 60  
 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNWKNMVEOMHEDIISLWDQSLKPCVK 120  
 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNWKNMVEOMHEDIISLWDQSLKPCVK 120  
 121 LTPCLVTLNCTDLRNTTNTNNSNSEGTIKGEMKNCSEFNATSIIGDRMQKEYAL 180  
 121 LTPCLVTLNCTDLRNTTNTNNSNSEGTIKGEMKNCSEFNATSIIGDRMQKEYAL 180  
 181 LYKLDIEPDNDNTSRLISCNTSVITQACPKISFEPPIHYCAPAGFALLKNDKFFSG 240  
 181 LYKLDIVSIDNDSTSRLISCNTSVITQACPKISFEPPIHYCAPAGFALLKNDKFFSG 240  
 241 KGSKCKNVSTVQCHGIRPVVSTOLLNGSLAEVWIRSENFNTNNAKTIIVQLREPVKN 300  
 241 KGSKCKNVSTVQCHGIRPVVSTOLLNGSLAEVWIRSENFNTNNAKTIIVHLNESVQIN 300  
 301 CSRPNNNTKSIIPWGPGRAPYTTGQIIGDIRQAHNCNISKNTNNAKLVQVVEKLGEOFNKT 360  
 301 CTFPNYKRXKRIHIGGRAPYTTGQIIGDIRQAHNCNISRKNDTLRQIVSKLKEQFNK 360  
 361 KIYFTNSSGGDPPIVTHSFNCAGEFFYNTQTQFDSIWNSENCTWMTIRGLNNTGRNDTI 420  
 361 TIVFNQSSGGDPPIVHVSFNCGGEFFYNTSPLFNSTWNG-NNTWN----NLTGSNNNI 414  
 421 TLPCKRIKOIINRQVEKAMVAPPIKGNISCSNITGLLITRDGGKDNDRDGNFRPG 480  
 415 TLOCKIKIINRQVEKAMVAPPIEQIRCSNITGLLITRDGGKDTDND-TEIFRPG 473

QY 481 GGDMDRNWSELYKYKVKIPLGVAPTAKRRVQRBBERAVGLGAMFFGFLGAAGSTMG 540  
 DB 474 GGDMDRNWSELYKYKVVVTIEPLGVAPTAKRRVQRBBERAVGLGAMFFGFLGAAGSTMG 532  
 QY 541 AASVTLTVQARQLLSGIVOOOSNLLRAIEAQOHLLOLTVMGIKOLQARILAVERYLKDQO 600  
 DB 533 AASVTLTVQARQLLSGIVOOONLLRAIEAQOHLLOLTVMGIKOLQARVLAVERYLKDQO 592  
 QY 601 LLGIMGCSGKLICTTTVPWNASMSKNTLEAIWNNMTWQWDEIDNYTSLIYSLIBESP 660  
 DB 593 LLGFWGCSGKLICTTTVPWNASMS-NKSLDDIWNNTWQWEREIDNYTSLIYSLLEKSQ 651  
 QY 661 IQOEKNEQELLELDKWANLWNNFDDISNWLWYIKIFIMIVGGLVGLRIVFVVLIVNVRQ 720  
 DB 652 TQOEKNEQELLELDKWSLWNNFDDITNWLWYIKIFIMIVGGLVGLRIVFAVLIVNVRQ 711  
 QY 721 GYSPLSFQTRLPAPRGPDRPEETEEESGDRDRSGLLVDGFLTLIWDLRSICLFSYHR 780  
 DB 712 GYSPLSLQTRPPVPRGPDRPEETEEESGDRDRSGRLVHGFLAIIWDLRSILFSYHH 771  
 QY 781 LRDLLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIATAEGTDRVI 840  
 DB 772 -RDLIIAARIVELLGRGWEVLKYWNLLQYWSQELKNSAVSLFNATAIATAEGTDRVI 830  
 QY 841 QVLQVRGALLHIPTRIQGLERALL 866  
 DB 831 EVLQVRGALLHIPTRIQGLERALL 856

RESULT 6  
 AAW88111  
 ID AAW88111 standard; protein; 856 AA.  
 XX  
 XX AAW88111;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 09-APR-1999 (first entry)  
 XX  
 DE Predicted amino acid sequence of the MN-PH1 envelope (env) protein.  
 XX  
 XX MN-PH1 clone; HIV-1; HIV-1 strain BA-L; envelope protein; vaccine;  
 KW immunotherapy; HIV infection; immunogen; HIV-1 diagnosis; env; ds.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 XX US5869313-A.  
 XX  
 PD 09-FEB-1999.  
 XX  
 XX 14-MAY-1996; 96US-00647714.  
 XX  
 XX 17-OCT-1990; 90US-00599491.  
 PR 25-FEB-1993; 93US-00022835.  
 PR 15-FEB-1995; 95US-00388809.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 XX Gallo RC, Franchini G, Reitz MS, Lori FC, Popovic M, Gartner S;  
 PI Markham PD;  
 XX  
 XX WPI; 1999-152779/13.  
 DR  
 XX DNA encoding env protein of the human immune deficiency virus isolate BA-  
 PT L - useful for producing protein for use in vaccines, as assay reagent  
 PT and to generate antibodies.  
 XX  
 XX Example 1; Fig 3A-C; 87pp; English.  
 XX  
 CC The present sequence represents the predicted amino acid sequence of the  
 CC MN-PH1 clone envelope (env) protein. MN-PH1 is a Human immunodeficiency  
 CC virus type 1 (HIV-1) clone. The specification also describes the env and  
 CC rev coding sequences of the HIV-1 strain BA-L (ATCC 40890). BA-L is more

CC typical of United States isolates of HIV-1 than previously known strains.  
CC Recombinant, complete env protein of the BA-L strain is used as a vaccine  
CC component and for immunotherapy of existing HIV infections, to detect HIV  
CC -specific antibodies, e.g. in donated blood, and as an immunogen to raise  
CC specific antibodies, for HIV-1 diagnosis. (Updated on 17-OCT-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 856 AA;

Query Match 86.0%; Score 3970; DB 2; Length 856;  
Best Local Similarity 86.7%; Pred. No. 4.5e-209;  
Matches 751; Conservative 36; Mismatches 69; Indels 10; Gaps 6;  
QY 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPWKEATTTLCASDAKAY 60  
DB 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPWKEATTTLCASDAKAY 60  
QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMKNMVEQMHEDIISLWDSLKPCVK 120  
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMKNMVEQMHEDIISLWDSLKPCVK 120  
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180  
DB 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180  
QY 181 LYKLDIEPDNDNTSYRLSCNTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240  
DB 181 LYKLDIVSDNDSTSYRLSCNTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240  
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVVRSENFNTNNAKTIIVQUREPVKIN 300  
DB 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVVRSENFNTNNAKTIIVHLNESVQIN 300  
QY 301 CSPNNNTRKSIWPGGRAPYTTGIIIGDIRAHCMISKTNTNALKOVVEKLGEOFNKT 360  
DB 301 CTSPNTRKRIHIGGRAPYTTGIIIGDIRAHCMISKTNTNALKOVVEKLGEOFNKT 360  
QY 361 KIVFTNSSGDDPEIVTHSFNCAGEFFYCNFTTQFDSIWNSENGTWNITRGLNNTGENDTI 420  
DB 361 TIVFNQSSGDDPEIVHVSFNCAGEFFYCNFTTQFDSIWNSENGTWNITRGLNNTGENDTI 420  
QY 421 TLPCTRKOIINRWQEVGKAMAPPIKGNITSCSNITGLLITRDGGKDNDRGNETFRPG 480  
DB 415 TLOCKIKOINRWQEVGKAMAPPIKGNITSCSNITGLLITRDGGKDTDND-TEIFRPG 473  
QY 481 GGDMDNRSELKYKVKIEPLGVAPTAKRRVQREERAVGLGAMFFGLGAAGSTWG 540  
DB 474 GGDMDNRSELKYKVKIEPLGVAPTAKRRVQREERAVGLGAMFFGLGAAGSTWG 532  
QY 541 AASVTITVQARQLLSGIVQQSNLLRAIEAQHLLQITVWGIKQLQARILAVERYLKDOQ 600  
DB 533 AASVTITVQARQLLSGIVQQSNLLRAIEAQHLLQITVWGIKQLQARILAVERYLKDOQ 592  
QY 601 LLGWCSCGKLICTTTPVWNASKKTLEAIWNNMTWQKDEIDNYTSLYSLEESP 660  
DB 593 LLGWCSCGKLICTTTPVWNASK-NKSLDDIWNNTWQKDEIDNYTSLYSLEESQ 651  
QY 661 IQEKNQEQLLEUDKWNALWNPDIENLWYIKIFIMIVGGLVGLRIVFVLISVNRVQ 720  
DB 652 TOEKNEQELLEUDKWNALWNPDIENLWYIKIFIMIVGGLVGLRIVFVLISVNRVQ 711  
QY 721 GYSPLSFQTLRPPRGPDRPEEIEEGGDRDRSGLLVGFTLLIWDLSRLSCLFSYHR 780  
DB 712 GYSPLSFQTLRPPRGPDRPEEIEEGGDRDRSGLLVGFTLLIWDLSRLSCLFSYHR 771  
QY 781 LRDLILLIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDRVI 840  
DB 772 -RDLLIAIARIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDRVI 830  
QY 841 QVLQVRGALLHIPTRIOGLERALL 866  
DB 831 EVLQVRGALLHIPTRIOGLERALL 856

RESULT 7  
ADP04371  
ID ADP04371 standard; protein; 857 AA.  
XX  
AC ADP04371;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Envelope protein gp120 from antibody sensitive HIV strain.  
XX  
KW anti-HIV; vaccine; human immunodeficiency virus type-1;  
KW envelope glycoprotein; cross-reactive immune response; gp41; CD4-binding;  
KW co-receptor-binding; outer domain; gp120.  
XX  
OS Human immunodeficiency virus 1.  
XX  
FN WO2004045495-A2.  
XX  
PD 03-JUN-2004.  
XX  
PF 12-MAY-2003; 2003WO-US014721.  
XX  
PR 10-MAY-2002; 2002US-0379052P.  
XX  
(JACK-) JACKSON FOUND HENRY M.  
XX  
Quinnan GV;  
XX  
WPI; 2004-431796/40.  
DR N-PSDB; ADP04370.  
XX  
Identifying an HIV-1 envelope protein that produces a cross-reactive  
PT immune response in a mammal, comprises introducing amino acid  
PT substitutions into gp120, and identifying substitutions that produce a  
PT cross-reactive immune response.  
XX  
Disclosure; SEQ ID NO 4; 68pp; English.  
XX  
The invention relates to a method of identifying a human immunodeficiency  
CC virus type-1 (HIV-1) envelope protein which produces a cross-reactive  
CC immune response following administration in a mammal by: (a) substituting  
CC one or more amino acids in or near the gp41, CD4-binding, co-receptor-  
CC binding or outer domains of gp120; and (b) identifying one or more amino  
CC acid substitutions in the domains that produce a cross-reactive immune  
CC response following administration in a mammal. The method is useful for  
CC identifying an HIV-1 envelope protein, which produces a cross-reactive  
CC immune response following administration in a mammal. Compositions  
CC comprising such HIV-1 envelope protein may be used as HIV subunit vaccine  
CC or immunogenic composition that can induce antibodies that neutralize HIV  
CC strains across different phenotypes and clades, or for the prevention or  
CC therapy of HIV-1 infection. This sequence corresponds to the gp120  
CC envelope glycoprotein from a neutralisation antibody-sensitive (T-cell  
CC line adapted) MN strain of HIV-1 (MN-P).  
XX  
SQ Sequence 857 AA;

Query Match 85.4%; Score 3941.5; DB 8; Length 857;  
Best Local Similarity 85.7%; Pred. No. 1.6e-207;  
Matches 742; Conservative 46; Mismatches 69; Indels 9; Gaps 5;  
QY 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPWKEATTTLCASDAKAY 60  
DB 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPWKEATTTLCASDAKAY 60  
QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMKNMVEQMHEDIISLWDSLKPCVK 120  
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMKNMVEQMHEDIISLWDSLKPCVK 120  
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180  
DB 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180



Qy	181	LYKLDIEPI	DNDNTSYRLIS	CVITQACPKIS	FEPIPIHYCAPAG	FAILKCN	DKKFSG	240																	
Db	181	LYKLDIEPI	DNDNTSHRLIS	CVITQACPKIS	FEPIPIHYCAPAG	FAILKCN	DKKFSG	240																	
Qy	241	KGSCKNVST	VOCTHGIRPV	WSTQOLLN	GLAAEEVVIR	SFNTFNNA	KTIIVOLRP	PKIN 300																	
Db	241	KGSCKNVST	VOCTHGIRPV	WSTQOLLN	GLAAEEVVIR	SFNTFNNA	KTIIVHLKESV	QVIN 300																	
Qy	301	CSRENNNTR	KSIPIMGPGRA	FYTTGGIIGD	IIRQAH	CNIISKTNN	TNALKQVVEK	LGQFNKT 360																	
Db	301	CTRPYNNK	RKRHIHGPGRA	FYTTKNIKGT	IIRQAH	CTISRAKWN	DTLRQIVSLK	QGFKNK 360																	
Qy	361	KIVFTNSG	GGDPEIVTHS	FNCA	GGEPFYNTTOL	FD	SIWNS	ENG	TWNIIRGLN	TGNDTI 420															
Db	361	TIVFNPSG	GGDPEIVMHS	FNCG	GGEPFYNTS	PLFN	TWNG	NTTN	W----	NTTGSNNNI 414															
Qy	421	TLPCRI	KQIINRMQ	BOVGKAM	YAPPTK	GNISCS	NITGLLL	TFD	GGKDDN	SRDGNTE	FRG 480														
Db	415	TQCKV	KQIINWQV	KAM	YAPPTK	GNISCS	NITGLLL	TFD	GGKDDN	SRDGNTE	FRG 473														
Qy	481	GGDMRD	NWRS	ELYKYK	VVIEPL	GVAPTKAK	RRVVQ	REERAV	GLGAM	FFGFLGA	AGSTMG 540														
Db	474	GGDMRD	NWRS	ELYKYK	VVIEPL	GVAPTKAK	RRVVQ	REERAV	GLGAM	FFGFLGA	AGSTMG 532														
Qy	541	AASVTL	TVQAR	QLLSG	IVQOQ	SNLLRA	TEAQO	HLQL	LTW	GWIKQ	LOARILAVE	RYLKDOQ 600													
Db	533	AASVTL	TVQAR	LLLSG	IVQOQ	SNLLRA	TEAQO	NMLQL	LTW	GWIKQ	LOARILAVE	RYLKDOQ 592													
Qy	601	LLGTW	CGSG	KLICTTT	VPWNAS	WSKNK	TEA	TWNN	TW	QDKE	IDNTY	TS	LIYS	LEESP 660											
Db	593	LLGFW	CGSG	KLICTTT	VPWNAS	WSKNK	TEA	TWNN	TW	QDKE	IDNTY	TS	LIYS	LEESP 651											
Qy	661	IQQEK	NEO	ELLELD	KWNL	WN	FDIS	NLWY	KIFIM	IVGG	I	VGLR	IVFV	VLIS	VNVR	VRQ 720									
Db	652	TQQEK	NEO	ELLELD	KWNL	WN	FDIS	NLWY	KIFIM	IVGG	I	VGLR	IVFV	VLIS	VNVR	VRQ 711									
Qy	721	GYSPL	SFQ	TRL	PAP	R	PD	R	PEIE	BEG	DR	DR	DR	SL	IVD	LR	SL	CF	SYHR 780						
Db	712	GYSPL	SFQ	TRL	PAP	R	PD	R	PEIE	BEG	DR	DR	DR	SL	IVD	LR	SL	CF	SYHR 771						
Qy	781	LRD	LL	IV	TR	IV	ELL	GR	GW	ELL	KY	WN	LL	QY	WS	QEL	KNS	AV	SL	NATA	IA	VAE	CTD	RV	840

[illegible]

Qy	481	G	D	M	R	D	N	R	S	E	L	Y	K	Y	K	V	K	I	E	P	L	G	V	A	P	T	K	A	K	R	V	V	Q	R	E	B	A	V	G	L	A	M	F	G	F	L	G	A	A	G	S	T	M	G	540			
Db	474	G	D	M	R	D	N	R	S	E	L	Y	K	Y	K	V	T	I	E	P	L	G	V	A	P	T	K	A	K	R	V	V	Q	R	E	K	R	A	-	A	I	G	A	L	F	L	G	F	L	G	A	A	G	S	T	M	G	532
Qy	541	A	A	S	V	T	L	T	V	Q	A	R	L	L	S	G	I	V	Q	Q	N	L	R	A	T	E	A	Q	H	L	L	Q	L	T	V	W	G	I	K	Q	L	A	R	L	I	A	V	E	R	Y	L	K	O	O	600			
Db	533	A	A	S	V	T	L	T	V	Q	A	R	L	L	S	G	I	V	Q	Q	N	L	R	A	T	E	A	Q	H	L	L	Q	L	T	V	W	G	I	K	Q	L	A	R	V	I	A	V	E	R	Y	L	K	O	O	592			
Qy	601	L	L	G	W	C	S	G	K	L	C	T	T	T	P	W	N	A	S	W	K	N	K	T	L	E	A	I	N	N	N	T	W	M	Q	D	K	E	I	D	N	T	S	L	I	S	L	T	E	S	P	660						
Db	593	L	L	G	W	C	S	G	K	L	C	T	T	T	P	W	N	A	S	W	S	-	N	K	S	D	D	I	N	N	N	T	W	M	Q	E	R	E	I	D	N	T	S	L	I	S	L	L	E	K	S	Q	651					
Qy	661	I	Q	E	K	N	E	O	E	L	E	L	D	K	W	A	N	L	N	N	F	D	I	S	N	M	L	Y	I	K	I	F	I	M	I	V	G	G	I	V	L	S	I	N	V	R	Q	720										
Db	652	T	Q	E	K	N	E	O	E	L	E	L	D	K	W	A	S	L	N	N	F	D	I	T	N	N	L	Y	I	K	I	F	I	M	I	V	G	G	I	V	L	S	I	N	V	R	Q	711										
Qy	721	G	Y	S	P	L	S	F	O	T	R	L	P	A	P	G	P	O	R	P	E	E	G	D	R	D	R	S	G	L	L	V	D	G	F	L	I	W	D	L	S	L	C	F	S	V	H	780										
Db	712	G	Y	S	P	L	S	L	T	R	P	P	V	P	G	P	O	R	P	E	E	G	D	R	D	T	S	G	R	L	V	H	G	F	L	I	W	D	L	S	L	F	L	S	V	H	771											
Qy	781	L	R	D	L	L	I	V	T	R	V	E	L	G	R	G	E	W	I	K	Y	W	N	L	L	Q	Y	S	O	E	L	K	N	S	A	V	S	L	F	N	A	T	A	I	A	V	A	E	G	T	D	R	V	I	840			
Db	772	-	R	D	L	L	I	A	A	R	I	V	E	L	G	R	G	E	W	I	K	Y	W	N	L	L	Q	Y	S	O	E	L	K	S	A	V	S	L	N	A	T	A	I	A	V	A	E	G	T	D	R	V	I	830				
Qy	841	Q	V	L	O	R	V	G	R	A	I	L	H	I	P	T	R	I	O	G	L	E	R	A	L	866																																
Db	831	E	V	L	O	R	A	G	R	A	I	L	H	I	P	T	R	I	O	G	L	E	R	A	L	856																																

RESULT 9

RESOL 3  
AAR79173

HAAR 19173  
TD AAR79173

XX  
ID AAK/91/3 standard; protein

XX  
AC  
79173.

AC  
AAR/91/3;  
VV

XX  
DE 00 MAR 1960

DT  
09-MAR-199

25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053

DE HIV virus-

**10**

KW HIV virus-

XX

OS Synthetic.

XX

**FH Key**

FT Protein

LE

FT Misc-diffe

### LE

Region

FT

Region

THE

Region

THE  
NOTES

Region

FT	FT	Region
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

XXIX

AA  
PN  
W09520162-

XX  
NA  
-29T02C60M

PT	gp120 peptide.
XX	
PS	Disclosure; Fig 1; 55pp; English.
XX	
CC	Peptide analogues of HIV virus antigen contain a plurality of peptide
CC	monomers, each comprising the gp120 CD4 binding region consensus sequence
CC	(amino acids 419-436) (see AAR79172). The synthetic peptomers are useful
CC	as therapeutic agents, immunogens in vaccine compositions or for the
CC	diagnosis of disease. Since the analogues maintain the conformation of
CC	the native proteins from which they are derived (the consensus sequence),
CC	they are useful against infection by various HIV-1 and HIV-2 isolates
CC	(see AAR79174-84)
XX	
SQ	Sequence 856 AA;

Query Match	85.2%	Score 3933;	DB 2;	Length 856;
Best Local Similarity	86.0%;	Pred. No. 4.7e-207;		
Matches	745;	Conservative 37;	Mismatches 74;	Indels 10; Gaps 6;
Qy	1	MRVKGIRRNQHWGWTMLGLGLMLCSATEKLUWTVVYGVPPVKCATTTFLFCASDAKAY	60	
Db	1	MRVKGIRRNQHWGWTMLGLGLMLCSATEKLUWTVVYGVPPVKCATTTFLFCASDAKAY	60	
Qy	61	DTEAHNWAATHACVPTDPNPQEVELNVNTEFNWKNMVEQMHEDIISLWDQSLKPCVK	120	
Db	61	DTEAHNWAATHACVPTDPNPQEVELNVNTEFNWKNMVEQMHEDIISLWDQSLKPCVK	120	
Qy	121	LTPLCVTLNCTDRLNTNTNNSNSSEGIKGEKMKCSFNTATISGDKMKQKEYAL	180	
Db	121	LTPLCVTLNCTDRLNTNTNNSNSSEGIKGEKMKCSFNTATISGDKMKQKEYAL	180	
Qy	181	LYKLDIPIINDNTSVRLISCNTSVITQACPKISFPIPIHYCAPAGFAILKCNDDKFKSG	240	
Db	181	LYKLDIPIINDNTSVRLISCNTSVITQACPKISFPIPIHYCAPAGFAILKCNDDKFKSG	240	
Qy	241	KGSKCNVSTVQCTHGIRPVVYSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLREPVKIN	300	
Db	241	KGSKCNVSTVQCTHGIRPVVYSTOLLNGSLAEVEVIRSENFNTNNAKTIIVHLNESVQIN	300	
Qy	301	CSPRNNNTRKSIIPWGPRAPIYTTQIIGDTRQAHCHISKTNTWNTALKQVVEKLEGEQPNKT	360	
Db	301	CTRPNYNKRKRRIHIGPRAPIYTTQIIGDTRQAHCHLSRSKWENTLKQIVTKLRVQPKNK	360	
Qy	361	KIVFTNSSGGDPELVTHSFNCAGEFFVYCNNTQLPDSIWSNENGTMNITRGLNNTGRNDTI	420	
Db	361	TIVFNRSGGDPELVHSMFCGGEFFVYCNNTPLFNSTWNG-NNTWN-----NTTGSNNNI	414	
Qy	421	TLPCRKQIINRWQEVGKAMYAPPKIGNITSNSSNITGLLLTRDGGKDNDRSGNETFRPG	480	
Db	415	TLQCKIKIINRWQEVGKAMYAPPKIGRCSNITGLLLTRDGGKDDTND-TEIFRPG	473	
Qy	481	GGDMRDNRSELKYKYVKVKEIPIGVAPTKAKRRVQVQREERAVGLGAMFFGFLGAAGTMG	540	
Db	474	GGDMRDNRSELKYKYVVTIPLGVAPTKAKRRVQVQREKRA-AIGALFLGFLGAAGTMG	532	
Qy	541	AASVTLTVQARQLLSGTVQOQSLLRAIEAQCHLLQLTWVGIIQOLOARIILAVERYLKDQO	600	
Db	533	AASVTLTVQARQLLSGTVQOQNNLLRAIEAQCHWMLQLTWVGIIQOLQARVLAVERYLKDQO	592	
Qy	601	LLGTWCGSGKLCITTTVPWNASWSKNKTLBAIWNNTMTWMDKEIDNYTSLIYSLIEESP	660	
Db	593	LLGFWCGSGKLCITTTVPWNASWS-NKSLDDIWNNTMTWQWEREIDNYTSLIYSLLEKSQ	651	
Qy	661	IQEKKEQELLEDKQWNLNWNFTDISNLWYIKIFIMIVGGLVGLRIVFVVLVIVNRVQ	720	
Db	652	TOQEKNEQELLEDKWSALNWNFTDNLWYIKIFIMIVGGLVGLRIVFVVLVIVNRVQ	711	
Qy	721	GYSPLSPOTPLAPRGDPREEEFGGDRDRSGLLVDGFTLTIVWDLRSILCLFSYHR	780	
Db	712	GYSPLSLQTPPVPRGDPREGIEEFGGERDRDTSGELVHGFLAIIVWDLRSILFLSYHH	771	
Qy	781	LRDLLLTVTRIVSLLGRGWIEILKYWNLLIQYWSQELKNSAVSLFNATAIAVAGTDRVI	840	





fractionating a crude gp120 preparation containing full-length, glycosylated gp120 using ion exchange chromatography so as to provide a first collection of fractions. A fraction from the first collection is selected that exhibits specific binding affinity for CD4 peptide, thereby producing a first fractionated material. The first fractionated material is fractionated by hydrophobic interaction chromatography so as to provide a second collection of fractions from which a second collection is selected that exhibits specific binding affinity for CD4 peptide. This second fraction is fractionated by size exclusion chromatography so as to provide a third collection of fractions exhibiting specific binding affinity for CD4 peptide, thereby providing the purified gp120. The purified gp120 can be used for antibody production and in vaccines. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 854 AA;

Query Match 84.8%; Score 3913; DB 2; Length 854;

Best Local Similarity 86.1%; Pred. No. 5.8e-206;

Matches 746; Conservative 35; Mismatches 73; Indels 12; Gaps 8;

QY 1 MRVKGIRRNQHWGCTMLGLLMICSAATEKLWTVVYGVVPMVKEATTLFCASDAKAY 60  
DB 1 MRVKGIRRNQHWGCTMLGLLMICSAATEKLWTVVYGVVPMVKEA-TTLFCASDAKAY 59  
QY 61 DTEAHNVWATHACVPTDPNPQVEVLNVNTEFNWKNMVEQMHEDIISLWDSLPFCVK 120  
DB 60 DTEAHNVWATHACVPTDPNPQVEVLNVNTEFNWKNMVEQMHEDIISLWDSLPFCVK 119  
QY 121 LTPCLVTLNCTDLRNTNTNNSNNSSEGTIKGEMKNCFSNATSGDKMKQKEYAL 180  
DB 120 LTPCLVTLNCTDLRNTNTNNSNNSSEGTIKGEMKNCFSNATSGDKMKQKEYAL 179  
QY 181 LYKLDIEPIDNDNTSVRLISNCTSVITQACPKISFEPIHYCAPAGFAILKCNCKKFSG 240  
DB 180 LYKLDIVSIDNSTSVRLISNCTSVITQACPKISFEPIHYCAPAGFAILKCNCKKFSG 239  
QY 241 KGCKXNVSTVQCTHGRPVVSTOLLNGSLAEBEVIRSENFTNNAKTIIVQLREPVKIN 300  
DB 240 KGCKXNVSTVQCTHGRPVVSTOLLNGSLAEBEVIRSEFTDNAKTIIVHLNESVQIN 299  
QY 301 CSRPNNTRKSIIPWGPRAEFTTGOIIGDTRQAHCNLSKTNWLNALQVVEKLGEOPNKT 360  
DB 300 CTRPNYKRRHIGPGRAYFTTNGIIGTRQAHCNISRAKNWATLQVSKUKEQFKNK 359  
QY 361 KIYFTNSSGGDPEIVTHSFNCAGEFFCYNTQLFDSIWNSENGTWNITRGLNNTGRNDTI 420  
DB 360 TIVFNQSSGGDPEIVMHSFNCGE-FYCNTPSLENSTWNGNSTWN-----NTTGSNNI 413  
QY 421 TLPCKIKQIINRQWGVKAMYPPIKGNISCSNITGLLTRDGGKDDNDRDGNETFRPG 480  
DB 414 TLQCKIKQIINRQWGVKAMYPPIEQIRCSNITGLLTRDGGKDDTND-TEIFRPG 472  
QY 481 GGMRDNRSELKYKVKLEPIGLVAPTAKRRVQREERAVGLGAMFFGLGAGSTMG 540  
DB 473 GGMRDNRSELKYKVKVTIEPLGVAPTAKRRVQREKKA-AIGALFLGLGAGSTMG 531  
QY 541 AASVTLTVQARQLLSGVVQOQSNLLRAIEAQHLLQLTWVGILQLOARILAVERYLKDDQ 600  
DB 532 AASVTLTVQARQLLSGVVQOQSNLLRAIEAQHLLQLTWVGILQLOARILAVERYLKDDQ 590  
QY 601 LLGIWCGSGKLICTTTVPWNASKNKTLBAIWNNTWMMQDKDEIDNTYSLIYSLIEESP 660  
DB 591 LLGIWCGSGKLICTTTVPWNASNS-KSLDDIWNNTWMMQWEREIDNTYSLIYSLIEKSQ 649  
QY 661 IQEKEQELLEDKQWNLNWNFDISNWLWYIKIFIMVGLVGLRIVFVVLIVNRVQ 720  
DB 650 TQEKNEQELLEDKQWNLNWNFDITNWLWYIKIFIMVGLVGLRIVFVVLIVNRVQ 709  
QY 721 GYSPLSFQTLPAKPRDPEEIEEGGDRDRDRSGLLVGGELTLIWDLRSLCLFSYHR 780  
DB 710 GYSPLSLQTRPPVPRGDRPEEGIEEGGDRDRDTSGLRVHGFLLAIWDLRSFLFSYH 769  
QY 781 LRDLILLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDRVI 840

DB 770 -RULLIARIVSELGRGWEVKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDRVI 828  
QY 841 QVLQVRGRALLHIPTIRQGLERALL 866  
DB 829 EVLQVRGRALLHIPTIRQGLERALL 854  
RESULT 13  
ABR55495  
ID ABR55495 standard; protein; 856 AA.  
XX ABR55495;  
XX 11-AUG-2003 (first entry)  
XX Amino acid sequence of a HIV envelope protein.  
XX Polymorphic marker; host response; HIV; AIDS related virus; vaccine;  
XX hepatitis related virus; HCV; HBV; drug resistance; envelope.  
XX Human immunodeficiency virus.  
XX WO2003035097-A1.  
XX 01-MAY-2003.  
XX 23-OCT-2002; 2002WO-AU001450.  
XX 23-OCT-2001; 2001AU-00008425.  
XX (BPIP-) BPIP PTY LTD.  
XX Mallal S;  
XX WPI; 2003-449231/42.  
XX

Determining the influence of variation in host genes on the selection of microorganisms with protein substitutions, comprises typing individuals of a cohort infected with a microorganism for an intrinsic polymorphic marker.

Claim 22; Page 90-91; 157pp; English.

The specification describes a method of determining the influence of variation in host genes on selection of microorganisms with protein substitutions. The method comprises typing all individuals of a population of patients infected with a microorganism for at least one selected intrinsic polymorphic marker involved in the host response to the presence of the microorganism. The method is useful for examining selective pressures confronting a wide range of organisms that exhibit pathogenic traits in a host, such as bacteria, fungi, mycobacterium, viruses and virus-like particles; for examining microorganisms that have adapted to evolve rapidly, including HIV and AIDS related viruses and the hepatitis related viruses such as HCV and HBV. The method is useful for designing a vaccine to prevent or delay the emergence of drug resistance in patients treated with a particular drug specific for a microorganism where the drug affects the replication of the microorganism at the nucleotide or amino acid level. The present sequence represents a HIV envelope protein, which is expected to provide optimal cytotoxic T lymphocyte (CTL) induced therapeutic protection to the cohort examined in that study

SQ Sequence 856 AA;

Query Match 84.4%; Score 3896; DB 6; Length 856;

Best Local Similarity 84.8%; Pred. No. 5e-205;

Matches 738; Conservative 45; Mismatches 69; Indels 18; Gaps 9;

QY 1 MRVKGIRRNQHW-WCGMTMLGLLMICSAATEKLWTVVYGVVPMVKEATTLFCASDAKAY 59  
DB 1 MRVKGIRRNQHW-WCGMTMLGLLMICSAATEKLWTVVYGVVPMVKEATTLFCASDAKAY 60

```
QY 60 YDTEAHNVWATHACVPTDPPNPQVEVLVNVTFENFMKNNVVEQMHEDIISLWDQSLKPCV 119
Db 61 YDTEHNVWATHACVPTDPPNPQVEVLVNVTFENFMKNNVVEQMHEDIISLWDQSLKPCV 120
QY 120 KLTPLCVTLNCTDLRNTNTNNSTDNNSSECTIKGEMKNCSEFNATISIGDKMOKEYA 179
Db 121 KLTPLCVTLNCTDLRNTNTNNSTDNNSSECTIKGEMKNCSEFNATISIGDKMOKEYA 174
QY 180 LLYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNDFKFS 239
Db 175 LFVKLDVVIDNDNTSYRLISCNSTSVITQACPKVSFEPPIHYCAPAGFAILKCNDFKKN 234
QY 240 GKSGCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVRISFNENFTNNAKTIIVOLRPVKI 299
Db 235 GTGCTNVSTVQCTHGIRPVVSTQLLNGSLAEEVVRISFNENFTNNAKTIIVQLNESVEI 294
QY 300 NCSRPNNTKSTP--MGCEAPYTTGQIIGDIRQAHCHNISKTNWNAKQVVEKLGEOF 357
Db 295 NCTRPNNNTKSTSIHIGPGRAFATGE--IGDIRQAHCHNISRABWNTLKQIVVKLREQF 353
QY 358 NKTGK-IVFTNSSGGDPEIVTHSFNCAGEFFYCNTTQLFDSIWNSENGTWNITRGLNNTGR 416
Db 354 GKNTKIVFQSSGGDPEIVHSHFNCGEFFYCNTTQLFNSFTWN--NSTWN--TESNNTG 410
QY 417 NDTITLPCRIRKQIINRWQEVGKAMYPPIKGNISCSNITGLLTRDGGKDDNSRDGNET 476
Db 411 NETITLPCRIRKQIINRWQEVGKAMYPPIRGQIRCSNITGLLTRDGGNNK---TET 467
QY 477 FRPGGDMRDNRSELYKVKVIEPLGVAPTAKERVVQREERAVGLGMFFGFLGAAG 536
Db 468 FRPGGDMRDNRSELYKVKVIEPLGVAPTAKERVVQREERAVGLGMFFGFLGAAG 527
QY 537 STWGAASVTLTVQARQLLSGIVQQNLLRAIEAQOHLQLTWGKIQOARILAVERYL 596
Db 528 STWGAASVTLTVQARQLLSGIVQQNLLRAIEAQOHLQLTWGKIQOARILAVERYL 587
QY 597 KQQLLGWCSGKLCITTTVPNWSKNTLEAIWNNTWQWQDEIDNTSLYSLI 656
Db 588 KQQLLGWCSGKLCITTAVPNTSWS--NKSINKIWDNNTWMEKEINNTYGTIYNLI 646
QY 657 EESPIQEKNEOELELDKWANLWDFDLSNWLWYIKIFIMIVGGVGLRIVFVLISVN 716
Db 647 EESQOQEKNEOELELDKWNLSWDFDLSKWLWYIKIFIMIVGGVGLRIVFVLISVN 706
QY 717 RVQGYSPLSFOTRLPAPRGPRDPEIEBEGDRDRDRSGLLVGDGLTLIWDRLSLCLF 776
Db 707 RVQGYSPLSFOTRLPAPRGPRDPEIEBEGDRDRDRSGLLVGDGLTLIWDRLSLCLF 766
QY 777 SYHRLDILLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAGT 836
Db 767 SYHRLDILLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAGT 826
QY 837 DRVIOVLORVGRALLHIPTRIOGLERALL 866
Db 827 DRIEVVQRACRAILHIPRIRQGVERRALL 856

RESULT 14
ADP04369
ID ADP04369 standard; protein; 863 AA.
AC
AC ADP04369;
XX
XX
DT 26-AUG-2004 (first entry)
XX
DE Envelope protein gp120 from antibody resistant HIV strain.
XX
KW anti-HIV; vaccine; human immunodeficiency virus type-1;
KW envelope glycoprotein; cross-reactive immune response; gp41; CD4-binding;
KW co-receptor-binding; outer domain; gp120.
XX
OS Human immunodeficiency virus 1.
XX
```

```
PN WO2004045495-A2.
XX 03-JUN-2004.
XX 12-MAY-2003; 2003WO-US014721.
XX 10-MAY-2002; 2002US-0379052P.
XX (JACK-) JACKSON FOUND HENRY M.
XX Quinman GV;
XX WPI; 2004-431796/40.
XX N-PSDB; ADP04368.
XX Identifying an HIV-1 envelope protein that produces a cross-reactive
PT immune response in a mammal, comprises introducing amino acid
PT substitutions into gp120, and identifying substitutions that produce a
PT cross-reactive immune response.
XX Disclosure; SEQ ID NO 2; 68pp; English.
XX The invention relates to a method of identifying a human immunodeficiency
CC virus type-1 (HIV-1) envelope protein which produces a cross-reactive
CC immune response following administration in a mammal by: (a) substituting
CC one or more amino acids in or near the gp41, CD4-binding, co-receptor-
CC binding or outer domains of gp120; and (b) identifying one or more amino
CC acid substitutions in the domains that produce a cross-reactive immune
CC response following administration in a mammal. The method is useful for
CC identifying an HIV-1 envelope protein, which produces a cross-reactive
CC immune response following administration in a mammal. Compositions
CC comprising such HIV-1 envelope protein may be used as HIV subunit vaccine
CC or immunogenic composition that can induce antibodies that neutralize HIV
CC strains across different phenotypes and clades, or for the prevention or
CC therapy of HIV-1 infection. This sequence corresponds to the gp120
CC envelope glycoprotein from a neutralisation antibody-resistant (primary)
CC MN strain of HIV-1 (MN-P).
XX Sequence 863 AA;
SQ
Query Match 83.6%; Score 3859.5; DB 8; Length 863;
Best Local Similarity 83.4%; Pred. No. 5e-203;
Matches 727; Conservative 47; Mismatches 83; Indels 15; Gaps 6;
QY 1 MRVKGIRRNQHWGWTMLLGLLMICSAATEKLWTVVYGVVPWKAEATTLFCASDAKAY 60
Db 1 MRVKGIRRNQHWGWTMLLGLLMICSAATEKLWTVVYGVVPWKAEATTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPPNPQVEVLVNVTFENFMKNNVVEQMHEDIISLWDQSLKPCVK 120
Db 61 DTEAHNVWATHACVPTDPPNPQVEVLVNVTFENFMKNNVVEQMHEDIISLWDQSLKPCVK 120
QY 121 LPLCVTLNCTDLRNTNTNNST-----DNNNSSEGTIKGEMKNCSEFNATISIGDKM 174
Db 121 LPLCVTLNCTDLRNTNTNNST-----DNNNSSEGTIKGEMKNCSEFNATISIGDKM 180
QY 175 QKEYALLYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCN 234
Db 181 QKEYALLYKLDIIVADKONTSYRLISCNSTSVITQACPKVSFEPPIHYCAPAGFAILKCN 240
QY 235 DKKPSGKSKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVRISFNENFTNNAKTIIVQLR 294
Db 241 DKNFTGKPGCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVRISFNENFTNNAKTIIVHLN 300
QY 295 EPVKINCSPPNNTRKSIPIWGPGRAFYTTGQIIGDIRQAHCHNISKTNWNAKQVVEKLG 354
Db 301 ESQVINCTRPYNNRRTRIHIGPGRAFYTTKNGIRQAHCTISSAKWNTLRQIVSKLK 360
QY 355 EGFNKTIVFTNSSGGDPEIVTHSFNCAGEFFYCNTTQLFDSIWNSENGTWNITRGLNNT 414
Db 361 EGFNKTIVFQSSGGDPEIVHSHFNCGEFFYCNTTSSLFNSTWNG--NNTWN----NTT 414
QY 415 GRNDITLPCRIRKQIINRWQEVGKAMYPPIKGNISCSNITGLLTRDGGKDDNSRDGN 474
```

Db 415 GSNSTITLOCKIQIINMMQEVGKAMYAPPIEQIRCSNITGLLTRDGNDDT-NNT 473  
Qy 475 ETRPFGGDMRDNWRSLEYKYKVKIEPLGVAPTAKRRVVQREERAVGLGAMFFGFLGA 534  
Db 474 ETRPFGGDMRDNWRSLEYKYKVKIEPLGVAPTAKRRVVQREKKA-AIGALFLGFLGA 532  
Qy 535 AGSTMGAASTVITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLQARILAVR 594  
Db 533 AGSTMGAASTVITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLQARILAVR 592  
Qy 595 YLKDQQLLGTWCGSKLICTTTPWNASKNTLEAINNMTMMQWDEIDNYSIYS 654  
Db 593 YLKDQQLLGTWCGSKLICTTTPWNASWS-NKSQEDIMNNMTMMQWERIDNYSIYE 651  
Qy 655 LIBESPTQEKNEQLELDKLANLWNFDSINLWYKIFIMVGLVGLRIVFVLSI 714  
Db 652 LLEKSNQKNEQLELDKLANLWNFDSINLWYKIFIMVGLVGLRIVFVLSI 711  
Qy 715 VNRVQGYSPSLQTRLPAPRGPDRPEETEEBEGDRDRDTSGLVDFLAIWVDRSL 774  
Db 712 VNRVQGYSPSLQTRLPAPRGPDRPEETEEBEGDRDRDTSGLVDFLAIWVDRSL 771  
Qy 775 LFSYHRLDLLIVTRIVELLGRGWELKYWNLLQYNSOELKNSAVSLFNATAVAE 834  
Db 772 LFSYHRLDLLIVTRIVELLGRGWELKYWNLLQYNSOELKNSAVSLFNATAVAE 831  
Qy 835 GTRDVITQVQRGRLHIIPTIRQGLERALL 866  
Db 832 GTRDVITQVQRGRLHIIPTIRQGLERALL 863

RESULT 15  
ADK14406  
ID ADK14406 standard; protein; 847 AA.  
XX  
AC ADK14406;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE HIV wild-type gp160 protein.  
XX  
KW HIV; gp120; gp41; glycoprotein; AIDS; gp140; variable loop;  
KW vaccine; viral disease; HIV infection; cytotoxic T lymphocytes;  
KW neutralising antibody.  
XX  
OS Human immunodeficiency virus 1; isolate JR-FL.  
XX  
PN US6710173-B1.  
XX  
PD 23-MAR-2004.  
XX  
PF 23-JUN-2000; 2000US-00602864.  
XX  
PR 25-JUN-1999; 99US-0141168P.  
XX  
PA (PROG-) PROGENICS PHARM INC.  
PA (ADAR-) ADARC AARON DIAMOND AIDS RES CENT.  
XX  
PI Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;  
XX WPI; 2004-256072/24.  
XX  
PT A nucleic acid encoding a modified viral env gene of an HIV-1 isolate,  
PT useful as a vaccine for treating or reducing the severity of a viral  
PT disease, e.g. HIV infection, or reducing the likelihood of a subject  
PT becoming infected with a virus.  
XX  
PS Disclosure; Page; 51pp; English.  
XX  
CC The invention relates to a nucleic acid encoding a modified viral  
CC envelope glycoprotein gene of an HIV-1 isolate comprising a nucleotide  
CC segment that encodes a modified form of an HIV-1 gp120 and gp41 complex,

CC where the modifications comprise a V35C, Y39C, W44C, P484C, G486C, A488C,  
CC F489C, T490C, or A492C mutation in gp120 and a D580C, W587C, T596C,  
CC V599C, P600C, or W601C mutation in gp41, such mutations being numbered by  
CC reference to the gp160 (precursor of gp41 and gp120) HIV-1 isolate JR-FL  
CC and the modifications resulting in a disulphide bond between gp120 and  
CC gp41 which stabilises the otherwise non-covalent gp120-gp41 interaction.  
CC The stabilization results in enhanced binding of the disulphide-bonded  
CC gp120-gp41 complex to HIV-1 neutralising antibodies and reduced binding  
CC of the complex to HIV-1 non-neutralising antibodies. Also included are a  
CC non-replicating viral vector comprising the nucleic acid cited above, a  
CC replicable vector comprising the nucleic acid, a host cell comprising the  
CC modified gp120 is further characterised by the absence of one or more of  
CC the variable loops present in the wild type gp120, these mutants being  
CC further described as gp140, a mutant of gp160 which has the proteolytic  
CC cleavage site between gp120 and gp41 mutated and a stop codon inserted  
CC prior to transmembrane domain of gp41 resulting in a 140kDa protein. The  
CC absent variable loop comprises V1, V2, V3 or its combination, preferably  
CC V1 and V2. The modified gp120 is further characterised by the absence or  
CC presence of one or more canonical glycosylation sites present in wild  
CC type gp120. One or more canonical glycosylation sites are absent from the  
CC V1V2 region of the gp120. The nucleic acids and proteins are useful as  
CC vaccines for treating a viral disease, e.g. HIV infection (AIDS),  
CC reducing the likelihood of a subject becoming infected with a virus,  
CC reducing the severity of a viral disease, or stimulating or enhancing in  
CC a subject production of antibodies or cytotoxic T lymphocytes. The  
CC present sequence is wild-type HIV gp160 protein. NOTE: This sequence is  
CC not shown in the specification but was obtained from genbank reference  
CC U63632 (DNA entry, proteins shown).  
XX  
SQ Sequence 847 AA;

Query Match 83.3%; Score 3845.5; DB 8; Length 847;  
Best Local Similarity 84.1%; Pred. No. 2.9e-202;  
Matches 728; Conservative 54; Mismatches 65; Indels 19; Gaps 6;  
Qy 1 MRVKGIRRYQHWHGWTMLLGLLMICSAATEKLWTVVYGVVPVWKEATTLFCASDAKAY 60  
Db 1 MRVKGIRKSYQYLWKGKGLLLGLIMTCSAVEKLWTVVYGVVPVWKEATTLFCASDAKAY 60  
Qy 61 DTEAHNVWATHACVPTDPNPQEVLEVNTENFNWKNVVEOMHEDIISLWDSLKPCVK 120  
Db 61 DTEAHNVWATHACVPTDPNPQEVLEVNTENFNWKNVVEOMHEDIISLWDSLKPCVK 120  
Qy 121 LTPCLVTLNCTDLRNTNTNNDNNNSSEGTIKGEMKNCFSNFIATISIDGMQKQYAL 180  
Db 121 LTPCLVTLNCKDV-NATNTN-----DSEGTMERGEIKNCFSNFIATISIDGMQKQYAL 172  
Qy 181 LYKLDIEPDNDNTSYRLISNTSVITQACPKISFEPPIHYCAPAGFAILKCNKDFSG 240  
Db 173 FYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPPIHYCAPAGFAILKCNKDFSG 232  
Qy 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
Db 233 KGPCKNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSDNFNTNNAKTIIVQLKESVEIN 292  
Qy 301 CSRPNNTNRKSIIPMGPRAFYTTGQIIGDIRQAHNCISKTNNALQVVKLGEQFNKT 360  
Db 293 CTRPNNTNRKSIHIGFGRFYTTGTEIIGDIRQAHNCISRAKWNDLKQIVIKLREQPENK 352  
Qy 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQLFDSIWNSENGTWNITRGLNTRNDTI 420  
Db 353 TIVFNHSSGGDPEIVMHSFNCGEFFYCNSTQLFNSNTN-----NNTGSENNTEG-TI 405  
Qy 421 TLPCLRKQIINWQEVGKAWYAPPIKNTSCSNITGLLTRDGGKDDNSRDNENFRPG 480  
Db 406 TLPCLRKQIINWQEVGKAWYAPPIKNTSCSNITGLLTRDGGKDDNSRDNENFRPG 462  
Qy 481 GDMRDNRSELYKYKVKVIEPLGVAPTAKRRVVQREERAVGLGAMFFGFLGAAGSTM 540  
Db 463 GDMRDNRSELYKYKVKVIEPLGVAPTAKRRVVQREERAVGLGAMFFGFLGAAGSTM 522  
Qy 541 AASVTITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLQARILAVERYLKDQ 600

Db	523	AAAMTLTVQARLLISGIVQQONLLRAIEAQRMQLQTVWGIKQARVLAVERYLGDOQ	582
Qy	601	LLGTWGCCKLICITTVPMNASKXKTLIAIWNNTWQWDKEIDNYTSLIYSLIEESP	660
Db	583	LLGIWGCCKLICITTVPMNASWS-NKSLDRIWNNTWMEWEREIDNYTSEIYTLIEESQ	641
Qy	661	IQOEKNEQELLELDKWANLWNFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISIVNRVQ	720
Db	642	NQOEKNEQELLELDKWASLWNFDTIKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRVQ	701
Qy	721	GYSPLSFQTLPPAPRGDPREEIEEGGDRDRDRSGLLVDGFTLIWVDLRSCLFSYHR	780
Db	702	GYSPLSFQTLPPAPRGDPREGIEEGGERDRDRSGRLVNGFTLAIWVDLRSCLFSYHR	761
Qy	781	LRDLLLIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI	840
Db	762	LRDLLLIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLNATAVAEGTDRII	821
Qy	841	QVLQVRGALLHIPTRIQGLERALL	866
Db	822	EALQTYTRAILHIPTRIQGLERALL	847

Search completed: April 11, 2005, 11:08:59  
Job time : 295.737 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:57:57 ; Search time 65.6201 Seconds  
(without alignments)  
1269.790 Million cell updates/sec

Title: US-09-762-261C-1

Perfect score: 4617

Sequence: 1 MRVKGIRRYQHWWGWTML.....GRALLHIPTRIQGLERALL 866

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79:\*

2: PIR1:\*

3: PIR2:\*

4: PIR3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3977.5	86.1	859	1 VCLJMN	env polyprotein pr
2	3862	83.6	852	2 T12016	envelope glycoprot
3	3853.5	83.5	847	2 S13289	env protein - huma
4	3845.5	83.3	847	2 T09448	envelope glycoprot
5	3839.5	83.2	843	1 H44001	env polyprotein pr
6	3775.5	81.8	861	1 VCLJSC	env polyprotein pr
7	3760	81.4	852	1 VCLJBR	env polyprotein -
8	3753.5	81.3	861	1 VCLJLV	env polyprotein pr
9	3744	81.1	856	1 VCLJH3	env polyprotein pr
10	3731	80.8	856	1 VCLJVL	env polyprotein pr
11	3716	80.5	854	2 S13288	env protein - huma
12	3710.5	80.4	851	2 S33985	env polyprotein -
13	3708	80.3	868	1 VCLJH4	env polyprotein pr
14	3703.5	80.2	861	1 VCLJKB	env polyprotein pr
15	3700	80.1	856	1 VCLJ3W	env polyprotein pr
16	3688.5	79.9	855	1 VCLJ42	env polyprotein pr
17	3523.5	76.3	859	2 T01672	envelope polyprote
18	3500.5	75.8	853	2 S54384	envelope polyprote
19	3500.5	75.8	855	1 VCLJZR	env polyprotein pr
20	3398	73.6	856	1 A44963	env polyprotein pr
21	3396	73.6	846	1 VCLJND	env polyprotein pr
22	3117	67.5	729	1 VCLJXK	env polyprotein pr
23	2746	59.5	854	1 VCLJ91	env polyprotein pr
24	2300.5	49.8	877	2 S49197	envelope protein p
25	2198	47.6	863	2 A50314	gag polyprotein -
26	2179	47.2	506	2 A40218	envelop glycoprote
27	1987.5	43.0	495	2 S31493	env polyprotein -
28	1963	42.5	443	2 C41621	env polyprotein p
29	1955	42.3	445	2 A41621	env polyprotein M

## ALIGNMENTS

### RESULT 1

VCLJMN

env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)  
N;Alternate names: coat polyprotein  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997  
C;Accession: A28922  
R;Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, B.; Farrell, K.; Wong-Staal  
Virology 164, 531-536, 1988  
A;Title: Envelope sequences of two new United States HIV-1 isolates.  
A;Reference number: A28922; MUID:88219542; PMID:3369091  
A;Accession: A28922  
A;Molecule type: DNA  
A;Residues: 1-859 <GUR>  
C;Genetics:  
C;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;30-859/Product: env polyprotein #status predicted <EPP>  
F;87,129,135,140,141,146,165,191,202,246,267,281,294,300,336,343,359,365,389,395,401,

30	1900.5	41.2	454	2	B41621	env polyprotein D
31	1669.5	36.2	357	2	S21994	envelope protein g
32	1661.5	36.0	357	2	S22006	envelope protein g
33	1648.5	35.7	357	2	S22004	envelope protein g
34	1644.5	35.6	357	2	S21996	envelope protein g
35	1636	35.4	358	2	S21998	envelope protein g
36	1629.5	35.3	357	2	S21992	envelope protein g
37	1580.5	34.2	357	2	S21990	envelope protein g
38	1579	34.2	358	2	S70417	envelope protein g
39	1574	34.1	358	2	S22000	envelope protein g
40	1568	34.0	358	2	S22002	envelope protein g
41	1368	29.6	858	1	VCLJG2	env polyprotein pr
42	1363	29.5	859	1	VCLJCT	env polyprotein pr
43	1355.5	29.4	885	2	S04322	env polyprotein -
44	1349	29.2	886	2	T11555	env protein - siml
45	1348.5	29.2	877	2	C46356	env polyprotein -

```
Db 361 TIVFNSSGGDPEIVHVSFNCGGEFFYCNTPSPFNSTWNG--NNTW--TKGNNTTSGNNNI 417
Qy 421 TLPCRILQIILNRQWGVCKANYAPIKGNISCSNITGLLLTRDGGKDDNSRDGNEFRPG 480
Db 418 TLOCKIQLIINMWQWGVCKANYAPIEGQIRCSNITGLLLTRDGGKDDTND--TEIFRPG 476
Qy 481 GGDMDNRWSELYKYKVVKIEPLGVAPTAKRRVVQREKRA-AIGALFLGFLGAAGSTWG 540
Db 477 GGDMDNRWSELYKYKVVVTEPLGVAPTAKRRVVQREKRA-AIGALFLGFLGAAGSTWG 535
Qy 541 AASVTTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQLTWVGIFKQLOARILAVERYLKDOQ 600
Db 536 AASVTTLTVQARQLLSGIVQOONLLRAIEAQHLLQLTWVGIFKQLOARILAVERYLKDOQ 595
Qy 601 LLGIWCSGKLICTTTVPWNASKNKTLKAIWNNNTWQWDEIDNYTSLIYSLIEEP 660
Db 596 LLGIWCSGKLICTTTVPWNASWS--NKSLEDDNNNTWQWDEIDNYTSLIYSLIEKSQ 654
Qy 661 IQQEKNEQELLELDKWANLWNFEDI SNLWYIKIFIMIVGGLVGLRIVFVLSIVNRVQ 720
Db 655 TQQEKNEQELLELDKWASLWNFEDI SNLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVQ 714
Qy 721 GYSPLSFQTRLPAPRGPDRPEIEEGGDRDRSGLLVGDFUTLWIDLRSLCLFSYHR 780
Db 715 GYSPLSFQTRLPAPRGPDRPEIEEGGDRDRSGLLVHGFALWIDLRSLCLFSYH 774
Qy 781 LRDLILIVTRIVELGRRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840
Db 775 -RDLILIAIVELGRRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 833
Qy 841 QVLQVRGALLHIPTRIQGLERALL 866
Db 834 EVLQVRGALLHIPTRIQGLERALL 859

RESULT 2
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: UNIPROT:O41883; EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g235
C:Genetics:
C:Superfamily: type B retrovirus env polyprotein

Query Match 83.6%; Score 3862; DB 2; Length 852;
Best Local Similarity 84.0%; Pred. No. 8e-269;
Matches 728; Conservative 54; Mismatches 69; Indels 16; Gaps 6;
Qy 1 MRVKGIRRNQHWGWTMLLGLLMTCATEKLWTVTVYGVVPWKATTLFCASDAKAY 60
Db 1 MRVKGIRKNCQHLWKWGTMLLGLLMTCASAEQLRVTVYGVVPWKATTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDPPQVELVNVTENFMNKNMVEQMHEDIISLWDSLPKVCVK 120
Db 61 DTEAHNVWATHACVPTDPPQVELVNVTENFMNKNMVEQMHEDIISLWDSLPKVCVK 120
Qy 121 LTPCLVTLNCTD-LRNTTNNSTDNNSSEGTIRKGEKNCSEFNATISIGDKMQKEYA 179
Db 121 LTPCLVTLNCTDYLNDTNT-----TSNNGGMEGGEIKNCSEFNATIRIGNKVQKEYA 173
Qy 180 LLYKLDIEFDNDNTSYRLISCNSTSVITQACPKISPEPIPIHYCTPAGFALLKCKDKFN 239
```

```
Db 174 LFYKLDVDPIDNTTTSYRLINCNTSVITQACPKVSEPIPIHYCTPAGFALLKCKDKFN 233
Qy 240 GKSGCKNVSTVQCTHGIRPVWSTQLLNLGSLAEVEVIRSENFNTNNAKTIIVQLRPPVKI 299
Db 234 GTGPTNVSTVQCTHGIRPVWSTQLLNLGSLAEVEVIRSENFNTNNAKTIIVQLNESVEI 293
Qy 300 NCSRPNNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCHNISKTNNWTNALKQVVEKLGQFNK 359
Db 294 NCTRPNNNTKRSIHIGPGQALYATGAIIGDIRQAHCHNISRAKWNNTLKQIVKLVQVQFN 353
Qy 360 TKIVFTNSGGDPEIVHVSFNCGGEFFYCNTPQPSIWNSENGTWNITRGLNNTGRNDT 419
Db 354 KTIIFNQSAGDPEIVHVSFNCGGEFFYCNTKLFNSTW--MFNNTWNTD---EDTEGNGT 409
Qy 420 ITLPCRILQIILNRQWGVCKANYAPIKGNISCSNITGLLLTRDGGKDDNSRDGNEFRP 479
Db 410 ITLPCRILQIILNRQWGVCKANYAPIRQIRCSNITGILLTRDGGTNNST---NETFRP 466
Qy 480 GGDMDNRWSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGLGAMPFGFGLGAAGSTM 539
Db 467 GGDMDNRWSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGLGFLGFLGAAGSTM 526
Qy 540 GAASVTTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQLTWVGIFKQLOARILAVERYLKQ 599
Db 527 GAASVTTLTVQARQLLSGIVQOONLLRAIEAQHLLQLTWVGIFKQLOARILAVERYLRDQ 586
Qy 600 QLLGIWCSGKLICTTTVPWNASWSKNKTLKAIWNNNTWQWDEIDNTSLIYSLIEES 659
Db 587 QLLGIWCSGKLICTTTVPWNASWS--NKSLEDKIWNNTWMEWEREIDNTSLIYSLIEES 645
Qy 660 PIQEKNEQELLELDKWANLWNFEDI SNLWYIKIFIMIVGGLVGLRIVFVLSIVNRVR 719
Db 646 QNQVKNQEQELLELDKWASLWSWFDITKWLWYIKIFIMIVGGLVGLRIVFVLSIVNRVR 705
Qy 720 QGYSPLSFQTRLPAPRGPDRPEIEEGGDRDRSGLLVGDFUTLWIDLRSLCLFSYH 779
Db 706 QGYSPLSFQTRLPAPRGPDRPEIEEGGDRDRSGLLVGDFUTLWIDLRSLCLFSYH 765
Qy 780 LRDLILIVTRIVELGRRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRV 839
Db 766 LRDLILIVARIVELGRRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRV 825
Qy 840 IQVLQVRGALLHIPTRIQGLERALL 866
Db 826 IEVQRTCAILHIPTRIQGLERALL 852

RESULT 3
S13289
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S13289
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13289
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <OBR>
A:Cross-references: UNIPROT:Q75760
C:Superfamily: type E retrovirus env polyprotein

Query Match 83.5%; Score 3853.5; DB 2; Length 847;
Best Local Similarity 84.2%; Pred. No. 3.2e-268;
Matches 729; Conservative 52; Mismatches 66; Indels 19; Gaps 6;
Qy 1 MRVKGIRRNQHWGWTMLLGLLMTCATEKLWTVTVYGVVPWKATTLFCASDAKAY 60
Db 1 MRVKGIRKSYQVWLKWKWGTMLLGLLMTCATEKLWTVTVYGVVPWKATTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDPPQVELVNVTENFMNKNMVEQMHEDIISLWDSLPKVCVK 120
```

```
Db 61 DTEHVNWATHACVPTDPNPQEVLENVTENFNMWKNMVQEQEDIISLDQSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNTSNDNNNSSEGIKKGEMKNCFSNFIATSIISLDQSLKPCVK 180
Db 121 LTPCLVTLNCKDV-NATNTN-----DSEGTMERGEIKNCFSNFIATSIISLDQSLKPCVK 172
QY 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFPIPIHYCAPAGFAILKCNCKKFSG 240
Db 173 FYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFPIPIHYCAPAGFAILKCNCKTFNG 232
QY 241 KGSCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLREPVKIN 300
Db 233 KGPCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLKESVEIN 292
QY 301 CSRPNNNTKRSIPMGPGRAFYTTGQIIGDIROAHNCISKTNTWNAKQVVEKLGEOFNKT 360
Db 293 CTRPNNNTKRSIHIGPGRAFYTTGQIIGDIROAHNCISRAKWNITLQIVIKLREQOFNK 352
QY 361 KIVFTNSSGDDPEIVTHSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db 353 TIVFNHSSGDDPEIVHVSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 405
QY 421 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
Db 406 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 462
QY 481 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAAGSTM 540
Db 463 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAAGSTM 522
QY 541 AASVTITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLKDQ 600
Db 523 AASMTITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLGDQ 582
QY 601 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660
Db 583 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 641
QY 721 GYSPLSFQTLAPRGPDRPEIEEGGDRDRSGLLVGDLFTLWDLRSICLPSYHR 780
Db 702 GYSPLSFQTLAPRGPDRPEIEEGGDRDRSGLLVGDLFTLWDLRSICLPSYHR 761
QY 781 LRDLILLITRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 840
Db 762 LRDLILLITRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 821
QY 841 QVLQVRGRALLHIPTRIQGLERALL 866
Db 822 EALQRTYRALLHIPTRIQGLERALL 847
```

## RESULT 4

```
T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T09448
R;Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
submitted to the EMBL Data Library, July 1996
A;Reference number: Z16673
A;Accession: T09448
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-847 <PAN>
A;Cross-references: UNIPROT:Q75760; EMBL:U63632; NID:gl465777; PID:gl465781
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
```

```
Query Match 83.3%; Score 3845.5; DB 2; Length 847;
Best Local Similarity 84.1%; Pred. No. 1.2e-267;
Matches 728; Conservative 54; Mismatches 65; Indels 19; Gaps 6;
QY 1 MRVKGIRRNQHWGWTMLLGLLMICSAATEKLWTVTVYGVVPWKEATTTLFCASDAKAY 60
Db 1 MRVKGIRKSYQYLWKGGLTLLGLLMICSAVEKLWTVTVYGVVPWKEATTTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQEVLENVTENFNMWKNMVQEQEDIISLDQSLKPCVK 120
Db 61 DTEAHNVWATHACVPTDPNPQEVLENVTENFNMWKNMVQEQEDIISLDQSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNTSNDNNNSSEGIKKGEMKNCFSNFIATSIISLDQSLKPCVK 180
Db 121 LTPCLVTLNCKDV-NATNTN-----DSEGTMERGEIKNCFSNFIATSIISLDQSLKPCVK 172
QY 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFPIPIHYCAPAGFAILKCNCKKFSG 240
Db 173 FYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFPIPIHYCAPAGFAILKCNCKTFNG 232
QY 241 KGSCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLREPVKIN 300
Db 233 KGPCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLKESVEIN 292
QY 301 CSRPNNNTKRSIPMGPGRAFYTTGQIIGDIROAHNCISKTNTWNAKQVVEKLGEOFNKT 360
Db 293 CTRPNNNTKRSIHIGPGRAFYTTGQIIGDIROAHNCISRAKWNITLQIVIKLREQOFNK 352
QY 361 KIVFTNSSGDDPEIVTHSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db 353 TIVFNHSSGDDPEIVHVSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 405
QY 421 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
Db 406 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 462
QY 481 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAAGSTM 540
Db 463 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAAGSTM 522
QY 541 AASVTITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLKDQ 600
Db 523 AASMTITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLGDQ 582
QY 601 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660
Db 583 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 641
QY 661 IQOEKNEQELLELDKQANLWNNWFDISNWLWYKIFIMIVGGLVGLRIVFVLSIVNRVQ 720
Db 642 NOOEKNEQELLELDKQANLWNNWFDITKWLWYKIFIMIVGGLVGLRIVFVLSIVNRVQ 701
QY 721 GYSPLSFQTLAPRGPDRPEIEEGGDRDRSGLLVGDLFTLWDLRSICLPSYHR 780
Db 702 GYSPLSFQTLAPRGPDRPEIEEGGDRDRSGLLVGDLFTLWDLRSICLPSYHR 761
QY 781 LRDLILLITRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 840
Db 762 LRDLILLITRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 821
QY 841 QVLQVRGRALLHIPTRIQGLERALL 866
Db 822 EALQRTYRALLHIPTRIQGLERALL 847
```

## RESULT 5

```
H44001
env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
```

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: H44001  
R;Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
J. Virol. 66, 6587-6600, 1992  
A;Title: Complete nucleotide sequence, genome organization, and biological properties of  
A;Reference number: A44001; MUID:93021387; PMID:1404605  
A;Accession: H44001  
A;Molecule type: DNA  
A;Residues: 1-843 <LTY>  
A;Cross-references: UNIPROT:P35961; GB:M33258  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;19-35/Region: hydrophobic  
F;30-489/Product: coat protein gp120 #status predicted <GP1>  
F;490-843/Product: coat protein gp41 #status predicted <GP2>  
F;499-515/Region: hydrophobic  
F;673-689/Region: hydrophobic  
F;738-755/Domain: transmembrane #status predicted <TMN>  
F;87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435  
Query Match 83.2%; Score 3839.5; DB 1; Length 843;  
Best Local Similarity 84.0%; Pred. No. 3.2e-267;  
Matches 728; Conservative 46; Mismatches 68; Indels 25; Gaps 7;  
QY 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
DB 1 MRATEIKNYQHLWKGTTGLGLMLCSAAEQWLWTVVYGVVPWKEATTLFCASDAKAY 60  
QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120  
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120  
QY 121 LPLCVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 180  
DB 121 LPLCVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 173  
QY 181 LYKLDIEPIDNDNTSVRLISCVTSVITQACPKISEPIPIHYCAPAGFAILKCNCKKFSG 240  
DB 174 FYNLDVVPI--DNASVRLISCVTSVITQACPKVSEPIPIHYCAPAGFAILKCNCKKFSG 231  
QY 241 KGSKNVSTVQCTHGRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
DB 232 TGPCTNVSTVQCTHGRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQLNESWIN 291  
QY 301 CSRPNNTKRSIPWGPGRFYTTGQIIGDIRQAHNCISKTNTNNAKTIIVQLREPVKIN 360  
DB 292 CTRPNNTKRSIPWGPGRFYTTGQIIGDIRQAHNCISKTNTNNAKTIIVQLREPVKIN 351  
QY 361 K-IVFTNSSGGDPEIVTHSFNCAGEFFVCNTTQLPDSIWNSENGTWNITRGLNNTGRND 419  
DB 352 KTIIFNPSSGGDPEIVTHSFNCAGEFFVCNTTQLPDSIWNSENGTWNITRGLNNTGRN 400  
QY 420 ITLPCRKIQIINRWQEVGKAMYPPIKGNISCSNITGLLTDGKDNRSRDNFTFRP 479  
DB 401 ITLPCRKIQIINRWQEVGKAMYPPIRGQIRCSNITGLLTDGKDNRSRDNFTFRP 457  
QY 480 GGGDMRDNRSSELYKYKVKVKEPLGVAPTAKERVVQREERAVGLGALFGFAGAGSTM 539  
DB 458 GGGDMRDNRSSELYKYKVKVKEPLGVAPTAKERVVQREERAVGLGALFGFAGAGSTM 517  
QY 540 GAASVTLTVOARQLLSGIVQOQNLLRAIEAQHLLQLTWGIGKQIQAARILAVERYLKQ 599  
DB 518 GAASVTLTVOARQLLSGIVQOQNLLRAIEAQHLLQLTWGIGKQIQAARILAVERYLKQ 577  
QY 600 QLLGIWCGSKLICCTTVPWNASNKNLEAIAWNNWQWQDKEDINTVTSIYLSIBES 659  
DB 578 QLLGIWCGSKLICCTTVPWNTSWS--NKSINEIWDNNTWKKWEREDINTVTHIYLSIQS 636  
QY 660 PLOQEKNEBELLELDKWNINWDFITKMLWYIKIFIMIVGGIIGLRIVFVLSIVNRV 719

DB 637 QNOQEKNEBELLALDKWASLWNNFDTKMLWYIKIFIMIVGGIIGLRIVFVLSIVNRV 696  
QY 720 QGVSPISFOTRLPAPGPPORPEIEEGGDRDRSGLLVGDLTLIWDLRSLCLPSYH 779  
DB 697 QGVSPISFOTRLPAPGPPORPDGIEEGGDRDRSGLLVGDLTLIWDLRSLCLPSYH 756  
QY 780 RLFDLLIVTRIVELLGRRGWEILKYWNNLQWVSOELKNSAVSLFNATAIAVAEGTDRV 839  
DB 757 RLFDLLIVTRIVELLGRRGWGLVWNNLQWVSOELKNSAVSLFNATAIAVAEGTDRV 816  
QY 840 IQVLQVRGRALLHIPRIQGLERALL 866  
DB 817 IEILQRAFRVFLHIPRIQGLERALL 843  
RESULT 6  
VCLJSC  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)  
N;Alternate names: coat polyprotein  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997  
C;Accession: B28922  
R;Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Staal  
Virology 164, 531-536, 1988  
A;Title: Envelope sequences of two new United States HIV-1 isolates.  
A;Reference number: A28922; MUID:88219542; PMID:3369091  
A;Accession: B28922  
A;Molecule type: DNA  
A;Residues: 1-861 <GUR>  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;30-861/Product: env polyprotein #status predicted <EPP>  
F;87,129,135,140,143,159,163,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396,  
Query Match 81.8%; Score 3775.5; DB 1; Length 861;  
Best Local Similarity 82.3%; Pred. No. 1.3e-262;  
Matches 716; Conservative 56; Mismatches 83; Indels 15; Gaps 9;  
QY 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
DB 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120  
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120  
QY 121 LPLCVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 180  
DB 121 LPLCVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 176  
QY 181 LYKLDIEPIDNDNTSVRLISCVTSVITQACPKISEPIPIHYCAPAGFAILKCNCKKFSG 240  
DB 177 FYKLDVVPIDNDTSTVTLNCTSVITQACPKVSEPIPIHYCA-RWFAILLNCKKFSG 235  
QY 241 KGSKNVSTVQCTHGRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
DB 236 TGPCTNVSTVQCTHGRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQLKEAVIN 295  
QY 301 CSRPNNTKRSIPWGPGRFYTTGQIIGDIRQAHNCISKTNTNNAKTIIVQLREPVKIN 358  
DB 296 CTRPNNTKRSIPWGPGRFYTTGQIIGDIRQAHNCISKTNTNNAKTIIVQLKEAVIN 355  
QY 359 KTKIIVFTNSSGGDPEIVTHSFNCAGEFFVCNTTQLPDSIWNSENGTWNITRGLNNTGRND 418  
DB 356 NKTIIIFNRSSGGDPEIVHSHFNCAGEFFVCNTTQLPDSIWNSENGTWNITRGLNNTGRND 411  
QY 419 TITLPCRKIQIINRWQEVGKAMYPPIKGNISCSNITGLLTDGKDNRSRDNFTFRP 476  
DB 412 TITLPCRKIQIINRWQEVGKAMYPPIKGNISCSNITGLLTDGKDNRSRDNFTFRP 471

233 FNGTGPCTNVSTVCTGTHIRPVPSTQLLNGSLAEVEVIRSENFNNVKTIIIVQLNESV 292  
298 KINGSRNNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCNISKNTNWTNALKQVVKLGEQF 357  
293 EINCTRENNTKRIITMGPGRVYVYTTGQIIGDIRRAHCNLSRSKWENTLKQIVTKLRVQF 352  
358 NKTKIVTNSGGDPEIVTHSFNCAGEFFYCNNTTOLFDSIWNSENCTWITRGLNNTGRN 417  
353 KNKTIVFNRSGGDPEIVMHSFNCGGEFFCNCNTTOLFENSTW-YRNTGTGNITEG-----N 405  
418 DTTPLCRIKQIINRWQEVGKAMYPPIKGNISCSNSNITGLLLTRDGGKDDNSRDGNETF 477  
406 SPITPLCRIKQIINRWQEVGKAMYPPIRGQIKCSSNITGLLLTRDGGNNNETTD-TEIF 464  
478 RPPGGDMRDNRWSLYKYKVKIIBPLGVAPTKAKRRVQVQBERAVGLGAMFFGLGAAGS 537  
465 RPPGGNMRDNRWSLYKYKVKIIBPLGVAPTKAKRRVQVQREKRAVGLGALFLGLGAAGS 524  
538 TMGAASVTLTVOARQLLSGIVQOQSNLLRATAEQHLLQLTWGIIKQLOARILAVERYLK 597  
525 TMGAASVTLTVOARQLLSGIVQOQNNLLMATEAQHMLLTWGIKQLOARVLAVERYLK 584  
598 DQQLLGWICSGKGLICTTVPWNASWSKNKLTLEAIWNNTNMTQWQKEIDNNTSLIYSLIE 657  
585 DQQLLGWICSGKGLICTTAVPWNASWS-NKSLSDIWDNMTWMEWEREIDNNTSLIYSLIE 643  
658 ESPIQEKNQELLEDKQWNLWNFDSINMLWYIKIFIMIVGGLVGLRIVFVLSIVNR 717  
644 DSQIQEKNKELELDKQWNLWNFNTINMLWYIKIFIMIVGGLTGLRIVFVLSIVNR 703  
718 VROQYSPLSFOTRPLPAPRGPPRPEIBEEGGDRDRSGLLVDGFLTLLIWDLSRSLCLFS 777  
704 VROQYSPLSFOTRPLPAPRGPPRPEIBEEGGDRDRSGPLVDGFLALFWDLRSFLFS 763  
778 YHRLRDLIIIVTRIVELLGRGWEILKYWMNLLQYWSQELKNSAVSLFNATAIABAGTD 837  
764 YHRLRDLIIIVTRIVELLGRGWEILKYWMNLLQYWSQELKNSAVSLLNATAIABGRTD 823  
838 RVIOVLQVRGPAHLIHIPRIROGLERAL 865  
824 RAIEWQVQRAFIILHIPRIROGLERAL 851  
  
RESULT 8  
VCLV1V  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-la)  
N;Alternate names: coat polyprotein  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 09-Jul-2004  
C;Accession: A03975  
R;Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
Cell 40, 9-17, 1985  
A;Title: Nucleotide sequence of the AIDS virus, LAV.  
A;Reference number: A90866; MUID:85099333; PMID:2981635  
A;Accession: A03975  
A;Molecule type: DNA  
A;Residues: 1-861 <WAI>  
A;Cross-references: UNIPROT:P03377; GB:K02013; NID:g326417; PIDN:AAB59751.1; PID:g326424  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein  
F;1-30/Domain: signal sequence #status predicted <SIG>  
F;31-516/Product: exterior membrane glycoprotein #status predicted <EXT>  
F;517-861/Product: transmembrane glycoprotein #status predicted <TMW>  
F;88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411,  
F;616,621,630,642,679,755,821/Binding site: carbohydrate (Agn) (covalent) #status predicted



```
Db 764 CLFSYHRLRDLIIIVRIVELLGRGWALKYWNLLQYWSQELKNSAVSLLNATAVA 823
Qy 834 EGTDRVIVQLVRGVRALLHIPTIRIQLERALL 866
Db 824 EGTDRVIEVVGAYRAIRHPRIRIQLERILL 856

RESULT 10
VCLJVL
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 09-Jul-2004
C;Accession: A03974
R;Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nature 313, 450-458, 1985
A;Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus
A;Reference number: A93355; MUID:85111157; PMID:2982104
A;Accession: A03974
A;Molecule type: DNA
A;Residues: 1-856 <MUE>
A;Cross-references: UNIPROT:P03376; GB:K02083; MID:g555008; PIDN:AAB59873.1; PID:g328559
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-511/Product: exterior membrane glycoprotein #status predicted <EXT>
F;512-856/Product: transmembrane glycoprotein #status predicted <TM>
F;88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406
F;611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 80.8%; Score 3731; DB 1; Length 856;
Best Local Similarity 82.1%; Pred. No. 2e-259;
Matches 717; Conservative 50; Mismatches 82; Indels 24; Gaps 11;

Qy 1 MRVKGIRRYQH---W-WGWTMLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 56
Db 1 MRVK---EKYQHLWRWGWKMGWMTLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 57
Qy 57 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFNMKNVWQVQHEHDIISLWQSLK 116
Db 58 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFNMKNVWQVQHEHDIISLWQSLK 117
Qy 117 PCVKLTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNSFNATSIGDKMQK 176
Db 118 PCVKLTPCLVSLKCTDLKNDTNT-----NSSSGRMIMKEGKNSFNISTSIGDKVQK 171
Qy 177 EYALLYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILLKCNCK 236
Db 172 EYAFFYKLDIIPIDNDTTSYRLISCNSTSVITQACPKVSPFPIPIHYCAPAGFAILLKCNCK 231
Qy 237 KPSGKSGCKNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSFENFNNAKTIIVQLREP 296
Db 232 TFGTGCTGVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSFENFNNAKTIIVQLNQS 291
Qy 297 VKINCRPNNTKRSIPM--GPGRAFVTTGQIIGDIRQAHCHNISKNTWTWALKQVVEKLG 354
Db 292 VEINCRPNNTKRSIRIQGPGRAFVTIGK--IGNMQAHCHNISRANKNTLQIADSKLR 350
Qy 355 EQFNKTK-IVFTNSSGGDPEIVTHSFNCAGEFFYCNTQLFDSIWNSENCTWNTIRGLNN 413
Db 351 EQFGNNTIIFKQSSGGDPEIVTHSFNCAGEFFYCNTQLFDSIWNSENCTWNTIRGLNN 407
Qy 414 TGRNDITLPCRQKQINRWQEVGKAMYPPIKGNISCSNITGLLTRDGGKDDNSRDG 473
Db 408 TEGSDITLPCRQKQINRWQEVGKAMYPPIKGNISCSNITGLLTRDGGKDDNSRDG 464
Qy 474 NETFRGGGDMRWNSSELYKVKVKEPLGVAPTAKRRVVOEERAVGLGAMFEGFLG 533
Db 465 SEIFRGGGDMRWNSSELYKVKVKEPLGVAPTAKRRVVOEERAVGLGAMFEGFLG 524
```

```
Qy 534 AAGSTWGAASVTLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTWGIKOLQARILAVE 593
Db 525 AAGSTWGAASMTLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTWGIKOLQARILAVE 584
Qy 594 RYLKDOQLLIGWCSGKLICTTTVPNNASWKNKTELEAIWNNMTWQMDKEIDNYSLIY 653
Db 585 RYLKDOQLLIGWCSGKLICTTTVPNNASW--NKSLEQIWNNTWMEWDREINNTSLIH 643
Qy 654 SLIEESPIQOEKNEQELLELDKWANLWNNPDISNWLWYIKIFIMIVGGLVGLRIVFVLS 713
Db 644 SLIEESQKQKNEQELLELDKWANLWNNLITNWLWYIKLFIWVGLVGLRIVFAVLS 703
Qy 714 INVRVQGYSPISFQTRLPAAPGPDPEETEEEGGDRDRDRSGLLVDGFTLTIWDRSL 773
Db 704 INVRVQGYSPISFQTLPTPRGDRPEGIEEGEDGERDRSIRLVNGSLALIWDDRSL 763
Qy 774 CLFSYHRLRDLIIIVRIVELLGRGWELLKYWNLLQYWSQELKNSAVSLLNATAVA 833
Db 764 CLFSYHRLRDLIIIVRIVELLGRGWALKYWNLLQYWSQELKNSAVSLLNATAVA 823
Qy 834 EGTDRVIVQLVRGVRALLHIPTIRIQLERALL 866
Db 824 EGTDRVIEVVGAYRAIRHPRIRIQLERILL 856

RESULT 11
S13288
env protein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S13288
R;O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A;Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c
A;Reference number: S13288; MUID:91043044; PMID:2172833
A;Accession: S13288
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-854 <OBR>
A;Cross-references: UNIPROT:Q85582; UNIPROT:Q72502; UNIPROT:Q90178; UNIPROT:Q78243; UNIP
C;Superfamily: type E retrovirus env polyprotein

Query Match 80.5%; Score 3716; DB 2; Length 854;
Best Local Similarity 81.9%; Pred. No. 2.4e-258;
Matches 715; Conservative 51; Mismatches 81; Indels 26; Gaps 12;

Qy 1 MRVKGIRRYQH---W-WGWTMLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 56
Db 1 MRVK---EKYQHLWRWGWKMGWMTLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 57
Qy 57 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFNMKNVWQVQHEHDIISLWQSLK 116
Db 58 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFNMKNVWQVQHEHDIISLWQSLK 117
Qy 117 PCVKLTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNSFNATSIGDKMQK 176
Db 118 PCVKLTPCLVSLKCTDLKNDTNT-----NSSSGRMIMKEGKNSFNISTSIGDKVQK 171
Qy 177 EYALLYKLDIEPIDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILLKCNCK 236
Db 172 EYAFFYKLDIIPVPI--DNTSYRLISCNSTSVITQACPKVSPFPIPIHYCAPAGFAILLKCNCK 229
Qy 237 KPSGKSGCKNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSFENFNNAKTIIVQLREP 296
Db 230 TFGTGCTGVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSFENFNNAKTIIVQLNQS 289
Qy 297 VKINCRPNNTKRSIPM--GPGRAFVTTGQIIGDIRQAHCHNISKNTWTWALKQVVEKLG 354
Db 290 VEINCRPNNTKRSIRIQGPGRAFVTIGK--IGNMQAHCHNISRANKNTLQIADSKLR 348
Qy 355 EQFNKTK-IVFTNSSGGDPEIVTHSFNCAGEFFYCNTQLFDSIWNSENCTWNTIRGLNN 413
```



Db 349 EQGNKNTIIFKQSSGGDEIVTHSFNCGGEFFVCNSTQLFNSTW---FNSTWS--TEGSSNN 405  
QY 414 TGRNDTITLPCRIOKIIINRQEVGKAMYAPPIKGNISCSNITGLLITRDGGKDDNSRDG 473  
Db 406 TEGSDTITLPCRIOKIIINRQEVGKAMYAPPIKGNISCSNITGLLITRDGG---NNNG 462  
QY 474 NETPRPGGDMRDNRSELYKYKVKIPIGLVAPTAKRRVVOREERAVGLGAMFFGFLG 533  
Db 463 SEIFRPGGDMRDNRSELYKYKVKIPIGLVAPTAKRRVVOREERAVGLGAMFFGFLG 522  
QY 534 AAGSTWGAASVTITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWVGIKQLOARILAVE 593  
Db 523 AAGSTWGAASVTITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWVGIKQLOARILAVE 582  
QY 594 RYLKDOQLLGGWCSGKLICTTTVPNNSKKNKTLKAIWNNMTWQWDEIDNYSILY 653  
Db 583 RYLKDOQLLGGWCSGKLICTTTVPNNSKKNKTLKAIWNNMTWQWDEIDNYSILY 641  
QY 654 SLIEESPIQOEKNEQELLELDKWNANLWNPFDISNWLWYIKIFIMIVGGLVGLRIVFVLS 713  
Db 642 SLIEESQOQEKNEQELLELDKWNANLWNPFDISNWLWYIKIFIMIVGGLVGLRIVFVLS 701  
QY 714 IVNRVQGYSPLSFQTHLPAPRGPDRPEEIEEGGDRDRSGLLVGDFTLTIWDLRL 773  
Db 702 IVNRVQGYSPLSFQTHLPAPRGPDRPEEIEEGGDRDRSGLLVGDFTLTIWDLRL 761  
QY 774 CLFSYHRLDLLIIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAI 833  
Db 762 CLFSYHRLDLLIIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAI 821  
QY 834 EGTDRVIVQVORVGRALLHIPRIROGLERALL 866  
Db 822 EGTDRVIVQVORVGRALLHIPRIROGLERALL 854  
  
RESULT 12  
S33985  
env polyprotein - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S33985  
R;Carlini, F.  
submitted to the EMBL Data Library, November 1991  
A;Reference number: S33979  
A;Accession: S33985  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-851 <CAR>  
A;Cross-references: UNIPROT:Q78243; EMBL:Z11530; NID:G60192; PIDN:CAA77628.1; PID:G60199  
C;Superfamily: type E retrovirus env polyprotein  
  
Query Match 80.4%; Score 3710.5; DB 2; Length 851;  
Best Local Similarity 81.7%; Pred. No. 5.9e-258;  
Matches 713; Conservative 51; Mismatches 80; Indels 29; Gaps 10;  
  
QY 1 MRVKGIRNTQH---W-WGWTMLLGLLMLTCSATEKLWTVVYGVVPWKAEATTLTFCASD 56  
Db 1 MRVKGIRNTQH---W-WGWTMLLGLLMLTCSATEKLWTVVYGVVPWKAEATTLTFCASD 57  
QY 57 AKAYDTEAHNVWATHACVPTDNPQEVVLNVTFENFNWKNVQWQWHDIIISLDQSLK 116  
Db 58 AKAYDTEAHNVWATHACVPTDNPQEVVLNVTFENFNWKNVQWQWHDIIISLDQSLK 117  
QY 117 PCVKLTPLCVLTLNCTDLRNTNTNNSDNNSEGTIKGEMKNSFNLIATSIGDKWQK 176  
Db 118 PCVKLTPLCVLTLNCTDLRNTNTNNSDNNSEGTIKGEMKNSFNLIATSIGDKWQK 171  
QY 177 EYALLKLDIPIEDNDNTSTRLISCNSTSVITQACPKISFPIPIHYCAPAGFPAIKCNCK 236  
Db 172 EYAFFYKLDIPIEDNDNTSTRLISCNSTSVITQACPKISFPIPIHYCAPAGFPAIKCNCK 231  
QY 237 KFGSKGCKNVSTVOCTHGIRPVVSTOLLNGSLABEEVIRSFNFNNAKTIIVOLREP 296

Db 232 TFNGTGPCTNVSTVQCTHGIRPVVSTOLLNGSLABEEVIRSFNFNNAKTIIVOLNTS 291  
QY 297 VKINCSRPNNNTKRSIPM--GPGRAFVTTCQIITCIRQAHCNISKTNWTNALQVVEKLG 354  
Db 292 VEINCTFPNNNTKRIQEGPGRAFVTIGK-IGNRQAHCNISRANKWNTLQKIDSKLR 350  
QY 355 EQFNKTK-IVFTNSSGGDEIVTHSFNCAGFFYCNNTQLFDSIWNSENCTWNTIGLNN 413  
Db 351 EQFGNNTIIFKQSSGGDEIVTHSFNCAGFFYCNNTQLFDSIWNSENCTWNTIGLNN 402  
QY 414 TGRNDTITLPCRIOKIIINRQEVGKAMYAPPIKGNISCSNITGLLITRDGGKDDNSRDG 473  
Db 403 TEGSDTITLPCRIOKIIINRQEVGKAMYAPPIKGNISCSNITGLLITRDGG---NSNN 459  
QY 474 NETPRPGGDMRDNRSELYKYKVKIPIGLVAPTAKRRVVOREERAVGLGAMFFGFLG 533  
Db 460 SEIFRPGGDMRDNRSELYKYKVKIPIGLVAPTAKRRVVOREERAVGLGAMFFGFLG 519  
QY 534 AAGSTWGAASVTITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWVGIKQLOARILAVE 593  
Db 520 AAGSTWGAASVTITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWVGIKQLOARILAVE 579  
QY 594 RYLKDOQLLGGWCSGKLICTTTVPNNSKKNKTLKAIWNNMTWQWDEIDNYSILY 653  
Db 580 RYLKDOQLLGGWCSGKLICTTTVPNNSKKNKTLKAIWNNMTWQWDEIDNYSILY 638  
QY 654 SLIEESPIQOEKNEQELLELDKWNANLWNPFDISNWLWYIKIFIMIVGGLVGLRIVFVLS 713  
Db 639 SLIEESQOQEKNEQELLELDKWNANLWNPFDISNWLWYIKIFIMIVGGLVGLRIVFVLS 698  
QY 714 IVNRVQGYSPLSFQTHLPAPRGPDRPEEIEEGGDRDRSGLLVGDFTLTIWDLRL 773  
Db 699 IVNRVQGYSPLSFQTHLPAPRGPDRPEEIEEGGDRDRSGLLVGDFTLTIWDLRL 758  
QY 774 CLFSYHRLDLLIIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAI 833  
Db 759 CLFSYHRLDLLIIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAI 818  
QY 834 EGTDRVIVQVORVGRALLHIPRIROGLERALL 866  
Db 819 EGTDRVIVQVORVGRALLHIPRIROGLERALL 851  
  
RESULT 13  
VCLJH4  
env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)  
N;Alternate names: coat polyprotein  
N;Contains: coat protein gp120; coat protein gp41  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C;Accession: C25523  
R;Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare, S.  
Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986  
A;Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human  
A;Reference number: A94136; MUID:87041461; PMID:3490666  
A;Accession: C25523  
A;Molecule type: DNA  
A;Residues: 1-868 <DES>  
A;Cross-references: UNIPROT:P05879; GB:M13137; NID:G326460; PIDN:AAA44311.1; PID:G326467  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei  
F;1-521/Product: coat protein gp120 #status predicted <GP1>  
F;522-868/Product: coat protein gp41 #status predicted <GP2>  
F;89,138,139,142,166,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459,  
  
Query Match 80.3%; Score 3708; DB 1; Length 868;  
Best Local Similarity 80.5%; Pred. No. 9.2e-258;  
Matches 703; Conservative 56; Mismatches 100; Indels 14; Gaps 7;  
  
QY 1 MRVKGIRNTQH---W-WGWTMLLGLLMLTCSATEKLWTVVYGVVPWKAEATTLTFCASD 60



Db 3 MRAKIGIRKCOHLRWGTMLLGMLMCSAAANLWVTYVYGVVWKEATTTLCFASDAKAY 62  
QY 61 DTAHNVWATHACVPTDNPQVELVNVNTEFNWKNMVEQMHEDIISLWDQSLKPCVK 120  
Db 63 DTAHNVWATHACVPTDNPQVELVNVNTEFNWKNMVEQMHEDIISLWDQSLKPCVK 122  
QY 121 LTPCLVTLNCTDLRNTTNTNNSDNNNSSEGTIKGEMKNCNFNATSIIGDQWQKEYAL 180  
Db 123 LTPCLVTLNCTDLRNTTNTNNSDNNNSSEGTIKGEMKNCNFNATSIIGDQWQKEYAL 181  
QY 181 LYKLDIEPIDN-----DNTSYRLISCNSTVITQACPKISPEPIPIHYCAPAGFAILKND 235  
Db 182 FYKLDVEPIDDKNTTNTKYLJNCNTSVITQACPKVSEPIPIHYCTPTGFALLKND 241  
QY 236 KPSGKSGCKNVSTVQCTHGRVWSTQLLNGSLAEVEVIRSENFNTNAKTIIVQVRE 295  
Db 242 KFNKGTPCTNVSTVQCTHGRVWSTQLLNGSLAEVEVIRSENFNTNAKTIIVQVNV 301  
QY 296 PVKINCSRPNNNTKSIIPMGPGAFYTTGQIIGDQWQKEYAL 355  
Db 302 SVBINCTRPNHNRKRVTLGPGRVWYTTGILNIRQAHNCNISRQNNNTLQOIATTLRE 361  
QY 356 QFNKTKIVFTNSGGDPEIVTHSFNCAGEFFYCNVTLQFDSIWN-SENGTWNITRLNNT 414  
Db 362 QFNKTKIAFNQSGDPEIVWHSFNCAGEFFYCNVTLQFDSIWN-SENGTWNITRLNNT 421  
QY 415 GRNDTITLPCRIKQIINRWQEVGKAMYAPPIKGNISCNSTVITQACPKISPEPIPIHYCAPAGFAI 474  
Db 422 G--DIITLPCRIKQIINRWQEVGKAMYAPPIKGNISCNSTVITQACPKISPEPIPIHYCAPAGFAI 476  
QY 475 ETRPGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVQVEERAVG-LGAMFPGFLG 533  
Db 477 ETRPGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVQVEERAVG-LGAMFPGFLG 536  
QY 534 AAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQIARILAVE 593  
Db 537 AAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQIARILAVE 596  
QY 594 RYLDQQLLGIWCGSKGLICTTTVPNNAWSKNKTLEAIWNNMTWQWKEIDNYSIY 653  
Db 597 RYLDQQLLGIWCGSKGLICTTTVPNNAWSKNKTLEAIWNNMTWQWKEIDNYSIY 655  
QY 654 SLIEEPIQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRVFWLS 713  
Db 656 TLIEEPIQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRVFWLS 715  
QY 714 INVRQGYSPLSFOTLPPAPRGDPEETEEBEGGDRDRSLGLVDFLTLWVDRSL 773  
Db 716 INVRQGYSPLSFOTLPPAPRGDPEETEEBEGGDRDRSLGLVDFLTLWVDRSL 775  
QY 774 CLFSYHRLDLLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIYA 833  
Db 776 CLFSYHRLDLLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIYA 835  
QY 834 EGTDRVIVLQVRGALLHPTRIQLERALL 866  
Db 836 EGTDRVIVLQVRGALLHPTRIQLERALL 868  
RESULT 14  
vcljkb  
env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp41)  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004  
C:Accession: A42995  
R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.  
Virology 189, 534-546, 1992  
A:title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated  
A:Reference number: A42995; MUID:92351552; PMID:1322587  
A:Accession: A42995

A:Molecule type: mRNA  
A:Residues: 1-861 <SHI>  
A:Cross-references: UNIPROT:P31819; GB:S41266; GB:D01206  
C:Genetics:  
A:Gene: env  
C:Superfamily: type B retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
P:1-689/Domain: extracellular #status predicted <EXT>  
F:1-33/Domain: signal sequence #status predicted <SIG>  
F:17-33/Region: hydrophobic #status predicted  
F:34-517/Product: coat protein gp120 #status predicted <CP1>  
F:518-861/Product: cleavage processing #status predicted  
F:518-861/Product: coat protein gp41 #status predicted <CP2>  
F:518-534/Region: hydrophobic #status predicted  
F:590-711/Domain: transmembrane #status predicted <TM1>  
F:712-861/Domain: intracellular #status predicted <INT>  
F:756-772/Region: hydrophobic #status predicted  
P:93,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414,  
Query Match 80.2%; Score 3703.5; DB 1; Length 861;  
Best Local Similarity 80.2%; Pred. No. 1.9e-257;  
Matches 704; Conservative 65; Mismatches 80; Indels 29; Gaps 12;  
QY 1 MRVKGIRRNQOH-W-WG-----WGTMLLGLLMICSAATEKLVVTVYGVVWKEATTTLCFCA 54  
Db 1 MRVKGIRRNQOH-W-WG-----WGTMLLGLLMICSAATEKLVVTVYGVVWKEATTTLCFCA 60  
QY 55 SDKAYDTBAHNVWATHACVPTDNPQVELVNVNTEFNWKNMVEQMHEDIISLWDQS 114  
Db 61 SDKAYDTBAHNVWATHACVPTDNPQVELVNVNTEFNWKNMVEQMHEDIISLWDQS 120  
QY 115 LKPCVKLTPCLVTLNCTDLRNTTNTNNSDNNNSSEGTIKGEMKNCNFNATSIIGDQW 174  
Db 121 LKPCVKLTPCLVTLNCTDLRNTTNTNNSDNNNSSEGTIKGEMKNCNFNATSIIGDQW 172  
QY 175 QKEYALLKLDIEPIDN-----NTSYRLISCNSTVITQACPKISPEPIPIHYCAPAGFAI 230  
Db 173 QKEYALLKLDIEPIDN-----NTSYRLISCNSTVITQACPKISPEPIPIHYCAPAGFAI 232  
QY 231 LKCNNDKFKSGKCKNVSTVQCTHGRVWSTQLLNGSLAEVEVIRSENFNTNAKTIIV 290  
Db 233 LKCNNDKFKSGKCKNVSTVQCTHGRVWSTQLLNGSLAEVEVIRSENFNTNAKTIIV 292  
QY 291 VQLREPVKINCSRPNNNTKSIIPMGPGAFYTTGQIIGDQWQKEYAL 350  
Db 293 VQLREPVKINCSRPNNNTKSIIPMGPGAFYTTGQIIGDQWQKEYAL 352  
QY 351 EKLGEQFNKTKIVFTNSGGDPEIVTHSFNCAGEFFYCNVTLQFDSIWNSENGTWNITRG 410  
Db 353 EKLGEQFNKTKIVFTNSGGDPEIVTHSFNCAGEFFYCNVTLQFDSIWNSENGTWNITRG 406  
QY 411 LNNTGRNDTITLPCRIKQIINRWQEVGKAMYAPPIKGNISCNSTVITQACPKISPEPIPIHYCAPAGFAI 468  
Db 407 WNGTG-PENITLPCRIKQIINRWQEVGKAMYAPPIKGNISCNSTVITQACPKISPEPIPIHYCAPAGFAI 465  
QY 469 NSRDGNETRPPGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVQVEERAVGLGAMF 528  
Db 466 NTSSSIEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVQVEERAVGLGAMF 525  
QY 529 FGFLGAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQIAR 588  
Db 526 FGFLGAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQIAR 585  
QY 589 ILAVERYLKQDQQLLGIWCGSKGLICTTTVPNNAWSKNKTLEAIWNNMTWQWKEIDNYSIY 648  
Db 586 ILAVERYLKQDQQLLGIWCGSKGLICTTTVPNNAWSKNKTLEAIWNNMTWQWKEIDNYSIY 644  
QY 649 TSIYSLIESPITQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRV 708  
Db 645 TSIYSLIESPITQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRV 704  
QY 709 FVVLISVNRVQGYSPLSFOTLPPAPRGDPEETEEBEGGDRDRSLGLVDFLTLWVDRSL 768  
Db 709 FVVLISVNRVQGYSPLSFOTLPPAPRGDPEETEEBEGGDRDRSLGLVDFLTLWVDRSL 768

Db 705 FTVLSIVNRVQGYSPLSFQTRLPA-RGDRPREGIEEGGERDRDRSGPLVDGLLAIWV 763  
QY 769 DLRLCLFSYHRLRDLILLIVTRIVELGRRGWELTKYWNLLQVWSQELKNSAVSLFNAT 828  
Db 764 DLRLCLFSYHRLRDLILLIVTRIVELGRRGWELTKYWNLLQVWSQELKNSAVSLFNAT 823  
QY 829 AIAVAEGTDRVIOVLQVRGRLHLLIPTRIRQGLERALL 866  
Db 824 AIAVAEGTDRVIELQRTYRAILLHIPVKIRQGLERALL 861

RESULT 15  
VCLJ3W  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 09-Jul-2004  
C:Accession: A24774  
R:Starcich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.  
Cell 45, 637-648, 1986  
A:Title: Identification and characterization of conserved and variable regions in the en  
A:Reference number: A24774; MUID:86218077; PMID:2423250  
A:Accession: A24774  
A:Molecule type: DNA  
A:Residues: 1-856 <STA>  
A:Cross-references: UNIPROT:P31872; GB:K03455; GB:M38432; NID:gi906382  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-501/Product: coat protein gp120 #status predicted <GP1>  
F:502-947/Product: coat protein gp41 #status predicted <GP2>  
F:87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,459

Query Match 80.1%; Score 3700; DB 1; Length 856;  
Best Local Similarity 81.5%; Pred. No. 3.4e-257;  
Matches 712; Conservative 49; Mismatches 87; Indels 26; Gaps 10;

QY 1 MRVKGIRRNQHWGCTMLLGLLMICSAATEKLMWTVYGVVWVKGATTTLFCSADAKAY 60  
Db 1 MRVKGIRRNQHLWIWGTMLFGWMMICSAVEQLWTVYGVVWVKGATTTLFCSADAKAY 60  
QY 61 DTEAHNVWATHACVPDPNPQVEVLNVVTENFNWKNMVEQMHEDIISLWDQSLKPCVK 120  
Db 61 STEAHKWATHACVPDPNPQVEVLNVVTENFNWKNMVEQMHEDIISLWDQSLKPCVK 120  
QY 121 LTPLCVTNLCTDLRNTNTNNSDNNSEGIKGMKNCSPNATSGDKMOKEYAL 180  
Db 121 LTPLCVTNLCTDLRNTNTNNSDNNSEGIKGMKNCSPNATSGDKMOKEYAL 170  
QY 181 LYKLDIEPIDNDN----TSYRLISNTSVITQACPKIPHYCAPAGFAILKND 235  
Db 171 FYKLDVVPKNSDSSYTRYRLHNTSVITQACSKVSEPIPIHYCAPAGFAILKND 230  
QY 236 KKFSGKSCKNVSTQCTHGIRPVWSTQLLNGSLAEEVIRSENFNTNNAKTIIVQLRE 295  
Db 231 KKFNGTGPCTNVSTQCTHGIRPVWSTQLLNGSLAEEVIRSENFNTNNAKTIIVHLNE 290  
QY 296 PVKINCSRPNNNT-RKSIIPNGPRAFTTQIIGDROAHNCISKYNWTNALKQVVEKLG 354  
Db 291 SVEINCTRPNNNRRHHIIGPRAFY-TGEIRGNTRQAHNCISRAKWNNTLKQIVEKLR 349  
QY 355 EQFNKTKIVFTNSGGDPEIVTHSFNCAGFFYCNNTTQLFDSIWNSENGTWNITRGLNNT 414  
Db 350 EQFNKTKIVFNHSSGGDPEIVTHSFNCGGFFYCDSTQLENSTWNV---TGISTEGNNT 406  
QY 415 GRN-DTITLPCRKIQRINQVQKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDG 473  
Db 407 EENGDTITLPCRKIQRINQVQKAMYPPIKGNISCSNITGLLLTRDGG---NSSSR 463  
QY 474 NETFRGGGDMRNWSELYKYKVKIEPLGVAPTAKRRVQREBRANG-LGAMPFGFL 532

Db 464 BEIFRPGGNMRDNWSELYKYKVKIEPLGVAPTAKRRVQREBRANGAIGAMFLGFL 523  
QY 533 GAAGSTMGAASVTLTYQARQLLSGIYQOOSNLIRAJEAQOHLIQLTWGKIQLOARILAV 592  
Db 524 GAAGSTMGAASVTLTYQARQLLSGIYQOQNNLLRAJEAQOHLIQLTWGKIQLOARILAV 593  
QY 593 ERYLKDQQLLGIWGCSSGKLICTTTVPWNASWSKNKTLEAIWNNMTWQMDKEIDNTYSLI 652  
Db 584 ERYLRDQQLLGIWGCSSGKLICTTTVPWNASWS-NKSMQDIWNNMTWMEWEREIDNTYSLI 642  
QY 653 YSLIEESPQOEKNEQELLEDKWNANWFDISNMLWYIKIPIMIVGGVGLRIVFVVL 712  
Db 643 YNLIBESQOEKNEQELLEDKWNANWFSITNMLWYIKIPIMIVGGVGLRIVFVSVL 702  
QY 713 SIVNRVQGYSPLSFQTRLPA-RGDRPREGIEEGGERDRDRSGPLVDGLTLLIWDLRS 772  
Db 703 SIVNRVQGYSPLSFQTHLPTPRGDPREGTEEGGERDRDRSVLVHGFALINWDLRS 762  
QY 773 LCLFSYHRLRDLILLIVTRIVELGRRGWELTKYWNLLQVWSQELKNSAVSLFNATAIAV 832  
Db 763 LCLFSYHRLRDLILLIVTRIVELGRRGWELTKYWNLLQVWSQELKNSAVSLFNATAIAV 822  
QY 833 AEGTDRVIOVLQVRGRLHLLIPTRIRQGLERALL 866  
Db 823 AEGTDRVIEVQVORICRAIIHIPRIHQGLERALL 856

Search completed: April 11, 2005, 11:15:28  
Job time : 77.6201 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:39:11 ; Search time 277.268 Seconds  
(without alignments)  
1599.394 Million cell updates/sec

Title: US-09-762-261C-1

Perfect score: 4617

Sequence: 1 MRVKGIRRYQHWWGWTML.....GRALLHIPRIQGLERALL 866

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4600	99.6	866	2 Q9WPZ4	Q9WPZ4 human immun
2	3990.5	86.4	861	2 Q90294	Q90294 human immun
3	3986.5	86.3	857	2 Q6BC02	Q6BC02 human immun
4	3970	86.0	856	1 ENV_HVIMN	P05877 human immun
5	3969.5	86.0	857	2 Q9YU15	Q9YU15 human immun
6	3964.5	85.9	861	2 Q90295	Q90295 human immun
7	3941.5	85.4	857	2 Q90296	Q90296 human immun
8	3931.5	85.2	853	2 Q7SVL6	Q7SVL6 human immun
9	3928	85.1	850	2 Q7SVL4	Q7SVL4 human immun
10	3917.5	84.8	855	2 Q7SVL0	Q7SVL0 human immun
11	3905	84.6	860	2 Q7SVL7	Q7SVL7 human immun
12	3896.5	84.4	863	2 Q92763	Q92763 human immun
13	3886	84.2	852	2 Q7SVL5	Q7SVL5 human immun
14	3886	84.2	854	2 Q40222	Q40222 human immun
15	3870.5	83.8	853	2 Q03811	Q03811 human immun
16	3868	83.8	865	2 Q8Q2X5	Q8Q2X5 human immun
17	3863.5	83.7	863	2 Q91U20	Q91U20 human immun
18	3862.5	83.7	863	2 Q91U29	Q91U29 human immun
19	3862	83.6	852	2 Q41883	Q41883 human immun
20	3859.5	83.6	863	2 Q8UNX2	Q8UNX2 human immun
21	3859	83.6	852	2 Q92761	Q92761 human immun
22	3857.5	83.5	863	2 Q91Y18	Q91Y18 human immun
23	3856.5	83.5	859	2 Q71260	Q71260 human immun
24	3854	83.5	866	2 Q91K01	Q91K01 human immun
25	3851.5	83.4	863	2 Q91Y17	Q91Y17 human immun
26	3848.5	83.4	859	2 Q7SVL2	Q7SVL2 human immun
27	3848	83.3	866	2 Q6EG58	Q6EG58 human immun
28	3848	83.3	866	2 Q6EK29	Q6EK29 human immun
29	3847.5	83.3	871	2 Q91K13	Q91K13 human immun
30	3846	83.3	858	2 Q6U239	Q6U239 human immun
31	3845.5	83.3	847	2 Q6BC19	Q6BC19 human immun

#### RESULT 1

Q9WPZ4 ID Q9WPZ4 PRELIMINARY; PRT; 866 AA.

AC Q9WPZ4;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Envelope protein.

OS Name=env;

OS Human immunodeficiency virus 1.

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

ON NCBI\_TaxID=11676;

OX [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99226722; PubMed-10221533; DOI=10.1089/088922299311088;

RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;

RT Expression and characterization of HIV type 1 envelope protein

associated with a broadly reactive neutralizing antibody response.;

RL AIDS Res. Hum. Retroviruses 15:561-570(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF128126; AAD40637.3; -

DR HSSP; P04578; 1DLB.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019028; C:viral capsid; IEA.

DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env GP41.

DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.

DR Pfam; PF00517; GP41; 1.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;

Transmembrane.

SQ SEQUENCE 866 AA; 98113 MW; 6E76021833F2EACD CRC64;

Query Match 99.6%; Score 4600; DB 2; Length 866;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 863; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRVKGIRRYQHWWGWTMLGLLMI CSATEKLWVTYVGVVVKATTTLCASDAKAY 60

Db 1 MRVKGIRRYQHWWGWTMLGLLMI CSATEKLWVTYVGVVVKATTTLCASDAKAY 60

QY 61 DTEAHNVWATHACVPTDPPQEVLVNVTENFNWKNMVEQMHEDIISLWDSLKPCVK 120

Db 61 DTEAHNVWATHACVPTDPPQEVLVNVTENFNWKNMVEQMHEDIISLWDSLKPCVK 120

QY 121 LTPCLCVTLNCTDLRNTNTNTNNNSSEGTIKGEMKNCNFNATSTGDKMQKEYAL 180

Db 121 LTPCLCVTLNCTDLRNTNTNTNNNSSEGTIKGEMKNCNFNATSTGDKMQKEYAL 180

QY 181 LYKLDIEPDNDNTSYRLISCNSTSVITQACPISFEPIPIHYCAPAGFAILKCNCKFSG 240

Db 181 LYKLDIEPDNDNTSYRLISCNSTSVITQACPISFEPIPIHYCAPAGFAILKCNCKFSG 240

32 3845.5 83.3 847 2 Q75760 human immun

33 3843 83.2 866 2 Q6EJY9 human immun

34 3842 83.2 850 2 Q6UYQ1 human immun

35 3840 83.2 870 2 Q8Q2X1 human immun

36 3839.5 83.2 843 1 ENV\_HVLY2 human immun

37 3838.5 83.1 853 2 Q6UYR2 human immun

38 3838.5 83.1 871 2 Q91K14 human immun

39 3838 83.1 866 2 Q6EK45 human immun

40 3834 83.0 866 2 Q6EK37 human immun

41 3833 83.0 852 2 Q6UZ34 human immun

42 3833 83.0 866 2 Q6EJZ7 human immun

43 3832.5 83.0 849 2 Q6UYQ5 human immun

44 3832 83.0 852 2 Q41885 human immun

45 3832 83.0 866 2 Q6EG67 human immun

#### ALIGNMENTS

Db 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240  
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
Db 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
QY 301 CSRPNNNTRKSIIPMGPGRAFYTTGQIIGDIRAHNCISKNTWTNALKQVVEKLGEQFNKT 360  
Db 301 CSRPNNNTRKSIIPMGPGRAFYTTGQIIGDIRAHNCISKNTWTNALKQVVEKLGEQFNKT 360  
QY 361 KIVFTNSSGGDPEIVTHSFNCAGEFFVNTQTFDSIWNSENGTWTITRGLNNTGRNDII 420  
Db 361 KIVFTNSSGGDPEIVTHSFNCAGEFFVNTQTFDSIWNSENGTWTITRGLNNTGRNDII 420  
QY 421 TLPCLRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480  
Db 421 TLPCLRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480  
QY 481 GGDMDNRWSELYKYKVVKIIEPLGVAPTAKRRVVQREERAVGLGAMFGFLGAAGSTWG 540  
Db 481 GGDMDNRWSELYKYKVVKIIEPLGVAPTAKRRVVQREERAVGLGAMFGFLGAAGSTWG 540  
QY 541 AASVTLTVOARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQLOARILAVERYLKDOQ 600  
Db 541 AASVTLTVOARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQLOARILAVERYLKDOQ 600  
QY 601 LLGIWCSGKLICTTIVPWNASKNKTLEAIWNNTWQWDEIDNYSLIYSLIEESP 660  
Db 601 LLGIWCSGKLICTTIVPWNASKNKTLEAIWNNTWQWDEIDNYSLIYSLIEESP 660  
QY 661 IQQEKNEQELLELDKWANLWNNFDSINLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720  
Db 661 IQQEKNEQELLELDKWANLWNNFDSINLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720  
QY 721 GYSPLSFQTLRPPAPGDRPEETEEBGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780  
Db 721 GYSPLSFQTLRPPAPGDRPEETEEBGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780  
QY 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840  
Db 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840  
QY 841 QVLQVRGALLHIPTIRIQGLERALL 866  
Db 841 EVLQVRGALLHIPTIRIQGLERALL 866

RESULT 2  
O90294 ID O90294 PRELIMINARY; PRT; 861 AA.  
AC O90294;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN Name=env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98362112; PubMed=9696803;  
RA Park E.J., Vujcic L.K., Anand R., Theodore T.S., Quinnan G.V. Jr.;  
RT "Mutations in both gp120 and gp41 are responsible for the broad  
RT neutralization resistance of variant human immunodeficiency virus type  
RT 1 MN to antibodies directed at V3 and non-V3 epitopes.";  
RL J. Virol. 72:7099-7107(1998).  
DR EMBL; AF075719; AAC33788.1; .  
DR HSP; P05877; INJ0  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
SQ SEQUENCE 861 AA; 97670 MW; 8136732E44445E56 CRC64;  
Query Match 86.4%; Score 3990.5; DB 2; Length 861;  
Best Local Similarity 86.8%; Pred. No. 8.2e-287;  
Matches 752; Conservative 38; Mismatches 67; Indels 9; Gaps 5;  
QY 1 MRVKGIRRYQHWWGWTMLLGLLMICSAATEKLWTVVYGVVPWKAEATTLFCASDAKAY 60  
Db 1 MRVKGIRRYQHWWGWTMLLGLLMICSAATEKLWTVVYGVVPWKAEATTLFCASDAKAY 60  
QY 61 DTEAHNVWATHACVTPDPNPQVELVNVTFENFMKNNVVEQMHEDIISLWQSLKPCVK 120  
Db 61 DTEAHNVWATHACVTPDPNPQVELVNVTFENFMKNNVVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPCLVTLNCTDLRNTNTNNSDNNNSSEGTIKGEMKNCNFNATISIGDKMQKEYAL 180  
Db 121 LTPCLVTLNCTDLRNTNTNNSDNNNSSEGTIKGEMKNCNFNATISIGDKMQKEYAL 180  
QY 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240  
Db 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240  
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
Db 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
QY 301 CSRPNNNTRKSIIPMGPGRAFYTTGQIIGDIRAHNCISKNTWTNALKQVVEKLGEQFNKT 360  
Db 301 CSRPNNNTRKSIIPMGPGRAFYTTGQIIGDIRAHNCISKNTWTNALKQVVEKLGEQFNKT 360  
QY 361 KIVFTNSSGGDPEIVTHSFNCAGEFFVNTQTFDSIWNSENGTWTITRGLNNTGRNDII 420  
Db 361 KIVFTNSSGGDPEIVTHSFNCAGEFFVNTQTFDSIWNSENGTWTITRGLNNTGRNDII 420  
QY 421 TLPCLRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480  
Db 421 TLPCLRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480  
QY 481 GGDMDNRWSELYKYKVVKIIEPLGVAPTAKRRVVQREERAVGLGAMFGFLGAAGSTWG 540  
Db 481 GGDMDNRWSELYKYKVVKIIEPLGVAPTAKRRVVQREERAVGLGAMFGFLGAAGSTWG 540  
QY 541 AASVTLTVOARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQLOARILAVERYLKDOQ 600  
Db 541 AASVTLTVOARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQLOARILAVERYLKDOQ 600  
QY 601 LLGIWCSGKLICTTIVPWNASKNKTLEAIWNNTWQWDEIDNYSLIYSLIEESP 660  
Db 601 LLGIWCSGKLICTTIVPWNASKNKTLEAIWNNTWQWDEIDNYSLIYSLIEESP 660  
QY 661 IQQEKNEQELLELDKWANLWNNFDSINLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720  
Db 661 IQQEKNEQELLELDKWANLWNNFDSINLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720  
QY 721 GYSPLSFQTLRPPAPGDRPEETEEBGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780  
Db 721 GYSPLSFQTLRPPAPGDRPEETEEBGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780  
QY 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840  
Db 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840  
QY 841 QVLQVRGALLHIPTIRIQGLERALL 866  
Db 841 EVLQVRGALLHIPTIRIQGLERALL 866



FT	DISULFID	125	201	By similarity.	
FT	DISULFID	130	162	By similarity.	
FT	DISULFID	223	252	By similarity.	
FT	DISULFID	233	244	By similarity.	
FT	DISULFID	301	335	By similarity.	
FT	DISULFID	381	445	By similarity.	
FT	DISULFID	388	418	By similarity.	
FT	CARBOHYD	87	87	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	129	129	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	135	135	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	140	140	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	141	141	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	146	146	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	161	161	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	165	165	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	191	191	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	202	202	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	246	246	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	267	267	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	281	281	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	294	294	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	300	300	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	336	336	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	343	343	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	359	359	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	365	365	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	395	395	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	401	401	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	405	405	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	406	406	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	413	413	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	448	448	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	465	465	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	612	612	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	617	617	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	626	626	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	638	638	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	816	816	N-linked (GLCNAC..)	(Potential).
SQ	SEQUENCE	856 AA;	97140 MW;	D197D809940BE732	CRC64;

Query Match 86.0%; Score 3970; DB 1; Length 856;  
Best Local Similarity 86.7%; Pred. No. 2.7e-285;  
Matches 751; Conservative 36; Mismatches 69; Indels 10; Gaps 6;

QY	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWTVVYGVVPWKATTLFCASDAKAY	60
Db	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWTVVYGVVPWKATTLFCASDAKAY	60
QY	61	DTBAHNVATHACVPTDPNPQVEYELVNVTFNFMWKNMVEQMHEDIISLWDQSLKPCVK	120
Db	61	DTEVHNWATQACVPTDPNPQVEYELVNVTFNFMWKNMVEQMHEDIISLWDQSLKPCVK	120
QY	121	LTPCLVTLNCTDLRNTNTNTNDNNSSEGTIKGEMKNCSPNATSGDKMKQKEYAL	180
Db	121	LTPCLVTLNCTDLRNTNTNTNANNNSSEGTIKGEMKNCSPNATSGDKMKQKEYAL	180
QY	181	LYKLDIETIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG	240
Db	181	LYKLDIVSDNDSTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG	240
QY	241	KGSKCNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQRLPEVKIN	300
Db	241	KGSKCNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVHLNESVQIN	300
QY	301	CSRPNNNTRKSIIPMGGRAPYTTGQIIGDIQROACHNISKNTWNTNALKQVVKLGEQFNKT	360
Db	301	CTRPNTNKRKRHIHGRCRAPYTTKNIIGTIQROACHNISRKNWNTLRQIVSKLGEQFNK	360
QY	361	KIVFTNSSGGDPIVTHSFNCAGEFFVCNTOLFDSIWNSENCTWNTIRGLNNTGNDTI	420
Db	361	TIVFNSSGGDPIVHVSFNCAGEFFVCNTSPFNSTWNG-NNTWN-----NTTGSNNNI	414
QY	421	TLPCRIRKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNFTFRPG	480

Db	415	TLQCKIKQIINRWQEVGKAMYPPIEGQIRCSNITGLLLTRDGGKDDTND-TEIFRPG	473
QY	481	GGDMRDNWESELYKYKVKIIEPLGVAPTAKERVVOREBVRVCLGAMFFGELCAAGSTWG	540
Db	474	GGDMRDNWESELYKYKVKIIEPLGVAPTAKERVVOREBVRVCLGAMFFGELCAAGSTWG	532
QY	541	AASVTITVQARQLLSGIVQQSNLLRAIEAQOHLQLTWGIKQLQARILAVERYLKDOQ	600
Db	533	AASVTITVQARQLLSGIVQQSNLLRAIEAQOHLQLTWGIKQLQARILAVERYLKDOQ	592
QY	601	LLGWCSCGKLICTTTPVWNASWKNKTLEAIWNNTWMMQDKIDNTYSLIYSLIEESP	660
Db	593	LLGFWCSCGKLICTTTPVWNASW-NKSLDDIWNNTWMMQDKIDNTYSLIYSLIEESQ	651
QY	661	IQOEKNEQELLELDKWNALWNNWFDISNMLWYIKIFIMIVGGLVGLRIVFVWLSIVNRVQ	720
Db	652	TOQEKNEQELLELDKWNALWNNWFDITNMLWYIKIFIMIVGGLVGLRIVFVWLSIVNRVQ	711
QY	721	GYSPISFOTRLPAPRGPDRPEIEBEGGDRDRSGLLVVDGFLTLIWDLSLCLFSYHR	780
Db	712	GYSPISLQTRFPVPRGPDRPEIEBEGGDRDTSGLVHGFLAIWVWDLRSLFLSYHH	771
QY	781	LRDLLIVTRIVELLGRRGWEILKYWNLLQYWSOELKNSAVSLFNATAIAVAEGTDRVI	840
Db	772	-RDLIIAARIVELLGRGWEVLKYWNLLQYWSOELKNSAVSLFNATAIAVAEGTDRVI	830
QY	841	OVLQVRGALLHIPTRIOGLERALL	866
Db	831	EVLQVRGALLHIPTRIOGLERALL	856

RESULT 5  
QYUL5 PRELIMINARY; PRT; 857 AA.  
AC QYUL5; (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN Name=env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98362112; PubMed=9696803;  
RA Park E.J., Vujcic L.K., Anand R., Theodore T.S., Quinnan G.V. Jr.;  
RT "Mutations in both gp120 and gp41 are responsible for the broad  
RT neutralization resistance of variant human immunodeficiency virus type  
RT 1 MN to antibodies directed at V3 and non-V3 epitopes.";  
RL J. Virol. 72:7099-7107(1998).  
DR EMBL; AF075722; AAC31819.1; -.  
DR HSSP; P05877; INJ0.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
SQ SEQUENCE 857 AA; 97095 MW; 922D362E10C60DA8 CRC64;

Query Match 86.0%; Score 3969.5; DB 2; Length 857;  
Best Local Similarity 86.4%; Pred. No. 2.9e-285;  
Matches 748; Conservative 41; Mismatches 68; Indels 9; Gaps 5;

QY	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWTVVYGVVPWKATTLFCASDAKAY	60
Db	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWTVVYGVVPWKATTLFCASDAKAY	60







Best Local Similarity 85.1%; Pred. No. 1.9e-282;  
Matches 739; Conservative 53; Mismatches 59; Indels 17; Gaps 6;

QY 1 MRVKGIRRYQHWWGWTMLLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFPCASDAKAY 60  
DB 1 MRVKGIRKYNQHLWRWGTMLLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFPCASDAKAY 60

QY 61 DTAHNWATHACVPTDNPQEVLPVNVTFNFMKNMVEQMHEDIISLWDSLPKPCVK 120  
DB 61 DTEVHNWATHACVPTDNPQEVLPVNVTFNFMKNMVEQMHEDIISLWDSLPKPCVK 120

QY 121 LTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNCNFNATSIIGDKMKEAYAL 180  
DB 121 LTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNCNFNATSIIGDKMKEAYAL 173

QY 181 LYKLDIEPDND--NTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKKF 238  
DB 174 FYKLDVVPIDNDNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKKF 233

QY 239 SGKSGCKNVSTVQCTHGIRPVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVK 298  
DB 234 SGKSGCKNVSTVQCTHGIRPVSTQLLNGSLAEVEVIRSDNFTNNAKTIIVQLREPVK 293

QY 299 INCSRPNNTKSIIPMGPGRAFYTTQIIGDIRQAHCNISKTWNTNALKOVVEKLGEQFN 358  
DB 294 INCTRPNNTRKSIHIGPGRAFYTTGIIIGDIRQAHCNLSRAKWDNTLKIIVKLRQFG 353

QY 359 KTKIVTNSGGDPEIIVTHSFNCAGFFYNTTQIIGDIRQAHCNLSRAKWDNTLKIIVKLRQFG 418  
DB 354 NKTIIVNQSGDPEIIVTHSFNCAGFFYNTTQIIGDIRQAHCNLSRAKWDNTLKIIVKLRQFG 418

QY 419 TITLPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDNDRDNETFR 478  
DB 407 TITLPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGG--NNESETEIFR 463

QY 479 PGGDMRDNRWSELYKYKVKIEPLGVAPTAKRRVVQREAVGLGAMPFGFLGAAGST 538  
DB 464 PGGDMRDNRWSELYKYKVKIEPLGVAPTAKRRVVQREAVGLGAMPFGFLGAAGST 523

QY 539 MGAASVTLTVQARQLLSGIVQQSNLLRLRAEAQHLQLTWTWGIKQIQAARILAVERYLKD 598  
DB 524 MGAASVTLTVQARQLLSGIVQQSNLLRLRAEAQHLQLTWTWGIKQIQAARILAVERYLKD 583

QY 599 QQLLGWCGSGKLICTTTPVWNASWKNKTLEAIWNNMTWMDKEIDNVTSLIYSLIEE 658  
DB 584 QQLLGWCGSGKLICTTTPVWNASWKNKTLEAIWNNMTWMDKEIDNVTSLIYSLIEE 642

RESULT 9

Q7SVL4

ID Q7SVL4

AC Q7SVL4

DT 01-OCT-2003 (T-EMBLrel. 25, Created)

DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)

DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

DE Envelope glycoprotein.

GN Name=env;

PRELIMINARY; PRT; 850 AA.

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

Q7SVL4

```

QY 659 SPIQEKNEQELLELDKWLNNWFDISNWLWYIKIFIMIVGGLVGLRIVFVLSIVNRV 718
Db 643 SQNQEKNEQELLELDKWLNNWFDITKWLWYIKIFIMIVGGLVGLRIVFVLSIVNRV 702
QY 719 RQGYSPLSFQTRLPAPRPPRPEIEBEGGDRDRSGLLVGFLTLIWDLSLCLFSY 778
Db 703 RQGYSPLSFQTRLPAPRPPRPEIEBEGGDRDRSGLLVGFLTLIWDLSLCLFSY 762
QY 779 HRLRDLILLIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 838
Db 763 HRLRDLILLIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 822
QY 839 VIOVLQVRGALLHIPTRIQGLERALL 866
Db 823 VIEVLQVTRAILHIPTRIQGLERALL 850

RESULT 10
Q7SVL0
ID Q7SVL0 PRELIMINARY; PRT; 855 AA.
AC Q7SVL0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
GN Names-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22628496; PubMed=12743293;
RA DOI=10.1128/JVI.77.11.6359-6366.2003;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
RA Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
RT "U.S. Human immunodeficiency virus type 1 epidemic: date of origin,
RT population history, and characterization of early strains.";
RL J. Virol. 77:6359-6366(2003).
DR ENBL; AY247225; AAP37153.1; -.
DR HSSP; P04578; IDLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00516; GP120; 1.
DR AIDS; C:coat protein; Envelope protein; Glycoprotein; Polyprotein;
DR Transmembrane.
KW Transmembrane.
SQ SEQUENCE 855 AA; 96890 MW; C373084C2DB3D06D CRC64;

Query Match 84.8%; Score 3917.5; DB 2; Length 855;
Best Local Similarity 85.3%; Pred. No. 2.1e-281;
Matches 740; Conservative 46; Mismatches 67; Indels 15; Gaps 6;

QY 1 MRVKGIRRNVOHWGCTMLGLLIMICSATEKLMVTVYVGVYVKEATTLFCASDAKAY 60
Db 1 MRVKGIRRNVOHLWRGTMGLMLCSAAEQLMVTVYVGVYVKEATTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQEVENVNFTFNWKNMVEQMHEDIISLWDSQSLPCVK 120
Db 61 DTEAHNVWATHACVPTDPNPQEVENVNFTFNWKNMVEQMHEDIISLWDSQSLPCVK 120
QY 121 LTPCLVTLNCTD-LRNTTNTNTNNDNNNSGRTIKGEMKNCNFNIATSIGDMQKEYA 179
Db 121 LTPCLVTLNCTDYLGNVTNTNTT-----SSGGTMEKGEIKGCSFNITTSIRDKVQKEYA 175
QY 180 LLYKLDIEPDNDNTSVRLISQNTSVITQCPKISEPIPIHYCAPAGFAILKCNKKPS 239
Db 176 LFYKLDVVPIDDDNTSVRLISQNTSVITQCPKISEPIPIHYCAPAGFAILKCNKKFN 235

```

```

QY 240 GKSGCKNVSTVQCTHGIRPVPVSTQLLNGSLAEEVVRSENFNTNNAKIIIVQLRPVKI 299
Db 236 GKSGCTNVSTVQCTHGIRPVPVSTQLLNGSLAEEVVRSENFNTNNAKIIIVQLNESVVI 295
QY 300 NCSRPNNTRKSIIPMGPGRAFYYTGTQIIGDIRAHCHNISKTNWTNALQVVEKLGQFNK 359
Db 296 NCTRPNNTRKSIHIGPGRAFYYTGTGIIIGNIROACHNLSRAEWNKLEQTVGKLRQFGN 355
QY 360 TKIVFTNSGGPPEIYVTHSPNCAGFEFFYCNTPOLFDSIWNSENGTWNITRGLNNTGRNDT 419
Db 356 KTIIVFNQSSGGPPEIYVTHSPNCAGFEFFYCNTPOLFSTW-----TWNTEGLNNTSEGN-- 408
QY 420 ITPLCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDN-SRDGNETFR 478
Db 409 ITPLCRIKQIINVMWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGNENGTEGTFR 468
QY 479 PGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVRQREERAVGLGAMFFGLGAAGST 538
Db 469 PGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVRQREKRAVGIGAVFLGFLGAAGST 528
QY 539 MGAASVTLTVQARQLLSGIVQQSNLLRAIEAQOHLQLTVMGIKOLOARILAVERYLKD 598
Db 529 MGAASITLTVQARLLLSGIVQQSNLLRAIEAQOHLQLTVMGIKOLOARILAVERYLKD 588
QY 599 QOLLGIWGGSGKLICTTTPVWNASWSKNTLEAIVNNMTWQMDKEIDNYSIYSLIEE 658
Db 589 QOLLGIWGGSGKLICTTAVPVNASWS-NKSLNEIWNMTWMEWEREINNYTGLIYNLIEE 647
QY 659 SPIQEKNEQELLELDKWLNNWFDISNWLWYIKIFIMIVGGLVGLRIVFVLSIVNRV 718
Db 648 SQNQEKNEQELLELDKWLNNWFDITNWLWYIKIFIMIVGGLVGLRIVFVLSIVNRV 707
QY 719 RQGYSPLSFQTRLPAPRPPRPEIEBEGGDRDRSGLLVGFLTLIWDLSLCLFSY 778
Db 708 RQGYSPLSFQTRLPAPRPPRPEIEBEGGDRDRSGLLVGFLTLIWDLSLCLFSY 767
QY 779 HRLRDLILLIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 838
Db 768 HRLRDLILLIVTRIVELLGRGWELIKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 827
QY 839 VIOVLQVRGALLHIPTRIQGLERALL 866
Db 828 VIEVLQVTRAILHIPTRIQGLERALL 855

RESULT 11
Q7SVL7
ID Q7SVL7 PRELIMINARY; PRT; 860 AA.
AC Q7SVL7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
GN Names-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22628496; PubMed=12743293;
RA DOI=10.1128/JVI.77.11.6359-6366.2003;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
RA Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
RT "U.S. Human immunodeficiency virus type 1 epidemic: date of origin,
RT population history, and characterization of early strains.";
RL J. Virol. 77:6359-6366(2003).
DR ENBL; AY247218; AAP37146.1; -.
DR HSSP; P20871; ICE4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.

```

DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
SQ SEQUENCE 860 AA; 97762 MW; A7E185F7BD421590 CRC64;  
  
Query Match 84.6%; Score 3905; DB 2; Length 860;  
Best Local Similarity 84.2%; Pred. No. 1.8e-280;  
Matches 734; Conservative 54; Mismatches 66; Indels 18; Gaps 7;  
  
QY 1 MRVKGIRRYQHWWGHTMLLGLLMICSAATEKLWVTYVYGVVPWKATTTTLFCASDAKAY 60  
DB 1 MRVKGIRKQHLWRGTMGLGLLMICSAAGKLWVTYVYGVVPWKATTTTLFCASDAKAY 60  
QY 61 DTEAHNVWATHACVPTDPNPQEVLELVNVTENFNNMKNMVEQMHEDIISLWDSQSLKPCVK 120  
DB 61 DTEIHNWATHACVPTDPNPQEVLELVNVTENFNNMKNMVEQMHEDIISLWDSQSLKPCVK 120  
QY 121 LTPCLVTLNCTDLRNTNTNNTNNSSEGTIKGEMKNCSEFNATSIATSGDKMKEAYAL 180  
DB 121 LTPCLVTLNCTDLRNTNTNNTNNSSEGTIKGEMKNCSEFNATSIATSGDKMKEAYAL 173  
QY 181 LYKLDIEPIDND--NTSYRLISNTSVITQACPKISPEPIPIHYCAPAGFAILKNDKXKF 238  
DB 174 FYRLDVVPIIDNDSTSYRLISNTSVITQACPKVSPPEPIPIHYCAPAGFAILKNDKXKF 233  
QY 239 SGKSGCKNVSTVQCTHGRVWSTQILLNGSLAEVEVIRSENFTNNAKTIIVQLREPVK 298  
DB 234 NGTGPCNTVSTVQCTHGRVPISTQILLNGSLAEVEVIRSDNFMNNAKTIIVQLNESVE 293  
QY 299 INCSRPNNTKSI PMGPGRAFTYTGIIIGIRQAHCNISKTNWTKALQVVEKLGEPFN 358  
DB 294 INCTRPNNTRKSIHIGPGRAFTYTGIIIGIRQAHCNLSRTKWNTRLQIVYKLRQFG 353  
QY 359 KTIIVFTNSGGDPEIVTHSFNCAGFEFFYCNLTQDFSIWNSB--NGTW-NITRGLNNT 414  
DB 354 NKTIVFNQSGGDPEIVTHSFNCAGFEFFYCDSTQFNSTNNDTLFNTNNDTRGSNNT 413  
QY 415 GRNDTITLPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLITRQGGKDDNSRDGN 474  
DB 414 --EDTITLPCRIKQIINRWQEVGKAMYAPPIRQIRCSNITGLLITRQGG--NNKSETT 469  
QY 475 ETRPGGDMRDNWRSELYKYKVKIPLGVAPTKAKRRVVRQREAVGLGAMPFGFLGA 534  
DB 470 EIPRPGGDMRDNWRSELYKYKVKIPLGVAPTKAKRRVVRQREAVGLGAMPFGFLGA 529  
QY 535 AGSTMGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQLOARILAVR 594  
DB 530 AGSTMGAASMTLTVQARQLLSGIVQOONLLRAIEAQHLLQITVWGIKQLOARILAVR 589  
QY 595 YLKDQQLLGIWCGSGKLICTTTPVWNASWKNKTLRAIWNMTWMDKEIDNYSILYS 654  
DB 590 YLKDQQLLGIWCGSGKLICTTTPVWNASW--NKSLDKIWNMTWMEWEINNTSLIYT 648  
QY 655 LIESPQKNEQLELLELDKMANLWNPDI SNLWYIKIFIMVGLGLRIVFVLSI 714  
DB 649 LIESQNEQKNEQLELLELDKMANLWNPDI TWLWYIKIFIMVGLGLRIVFVLSI 708  
QY 715 VNRVROGYSPLSPQTRLPAPRGPDRPEIEEGGDRDRDSGLLVDFLTLIWDLRSLC 774  
DB 709 VNRVROGYSPLSPQTRLPAPRGPDRPEIEEGGDRDRDSGLLVDFLTLIWDLRSLC 768  
QY 775 LFSYHRLDLLLIVTRIVELLGRGWEILKYWNLLQYMSQELKNSAVSLFNATAVAE 834  
DB 769 LFSYHRLDLLLIVTRIVELLGRGWEILKYWNLLQYMSQELKNSAVSLFNATAVAE 828  
QY 835 GTDRVIOVLQVRGALLHIPTIRIQGLERALL 866  
DB 829 GTDRIIIVLQVRVRAIIHIPTIRIQGLERALL 860

RESULT 12

092763 PRELIMINARY; PRT; 863 AA.  
ID 092763;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN Name-env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98178716; PubMed=9519894;  
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,  
RT Sheppard W.H.;  
RT "Diversity of the human immunodeficiency virus type 1 envelope  
RT glycoprotein in San Francisco Men's Health Study participants.";  
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).  
DR EMBL; AF025755; AAC40592.1; -;  
DR HSSP; P04578; 1DLB.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
SQ SEQUENCE 863 AA; 97863 MW; 13A9530DBBAEF4C CRC64;  
  
Query Match 84.4%; Score 3896.5; DB 2; Length 863;  
Best Local Similarity 84.3%; Pred. No. 7.7e-280;  
Matches 734; Conservative 49; Mismatches 75; Indels 13; Gaps 6;  
  
QY 1 MRVKGIRRYQHWWGHTMLLGLLMICSAATEKLWVTYVYGVVPWKATTTTLFCASDAKAY 60  
DB 1 MRVKGIRKQYHFWRGTMGLGLLMICSAAEKLWVTYVYGVVPWKATTTTLFCASDAKAY 60  
QY 61 DTEAHNVWATHACVPTDPNPQEVLELVNVTENFNNMKNMVEQMHEDIISLWDSQSLKPCVK 120  
DB 61 DTEAHNVWATHACVPTDPNPQEVLELVNVTENFNNMKNMVEQMHEDIISLWDSQSLKPCVK 120  
QY 121 LTPCLVTLNCTDLRNTNTNNTNNSSEGTIKGEMKNCSEFNATSIATSGDKMKEAYAL 180  
DB 121 LTPCLVTLNCTDLRNTNT--TANTSSSGGAMEEGDMKNCSEFNITTSIRNMQRERVL 177  
QY 181 LYKLDIEPIDNDNTSVRLISNTSVITQACPKISPEPIPIHYCAPAGFAILKNDKXKFSG 240  
DB 178 FYKLDIVPIDNDNTSVRLISNTSVITQACPKVSPPEPIPIHYCAPAGFAILKNDKXKFSG 237  
QY 241 KSGCKNVSTVQCTHGRVWSTQILLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300  
DB 238 TGPCENVSTVQCTHGRVWSTQILLNGSLAEVEVIRSENFTNNAKTIIVQLNASVEIN 297  
QY 301 CSRPNNTKSI PMGPGRAFTYTGIIIGIRQAHCNISKTNWTKALQVVEKLGEPFNK 359  
DB 298 CTRPNNTKSIHIGPGRAFTYTGIIIGIRQAHCNLSREKWTNLTQIVKELGEPFNK 357  
QY 360 KTIIVFTNSGGDPEIVTHSFNCAGFEFFYCNLTQDFSIWNSB--NGTW-NITRGLNNT 419  
DB 358 TTIIVFNQSGGDPEIVTHSFNCAGFEFFYCNLTQDFSIWNSB--NGTW-NITRGLNNT 413  
QY 420 ITPPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLITRQGGKDDNSRDGN--E 475  
DB 414 ITPPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLITRQGGKDDNSRDGN--E 473  
QY 476 TFRPGGDMRDNWRSELYKYKVKIPLGVAPTKAKRRVVRQREAVGLGAMPFGFLGA 535  
DB 474 IFRPGGDMRDNWRSELYKYKVKIPLGVAPTKAKRRVVRQREAVGLGAMPFGFLGA 533

```
QY 536 GSTGGAASVTLTVQAROLLGSIIVQOQNLLRAEAQOHLQLTVMGKIQARILAVERY 595
DB 534 GSTGGAASITLTVQAROLLGSIIVQOQNLLRAEAQOHLQLTVMGKIQARILAVERY 593
QY 596 LKQOQLLGIWGCCKLICTTTPVWNASWKNKTLKLEAIWNNWMTWQWDKEIDNTSLIYSL 655
DB 594 LVDQQLLGIWGCCKLICTTAVPWNASWS-NKGFKEIWNNTWWEKEINNTYRLIYTL 652
QY 656 IEESPIQOEKNEBELLELDKWNILWNNWFDISNWLWYIKIFIMIVGGLVGLRIVFVLSIV 715
DB 653 IEESQOEKNEBELLELDKWNILWNNWFDISNWLWYIRIFIMIVGGLVGLRIVFVLSIV 712
QY 716 NRVQGYSPLSFQTRLPAPGPRDPEIEEGDRDRDRSGLLVGFLTLIWDLSLCL 775
DB 713 NRVQGYSPLSLQTRPPAPGPRDPEIEEGDRDRDRSGLLVGFLTLIWDLSLCL 772
QY 776 FSHRLRDLILLIYTRIVELLGRGWELIKYWNLLQYWSQELKNSAVLSFNATAIAVAEG 835
DB 773 FSHRLRDLILLIYTRIVELLGRGWELIKYWNLLQYWSQELKNSAVLSFNATAIAVAEG 832
QY 836 TDRVIOVLQVRGRALLHIPRIROGLERALL 866
DB 833 TDRILIEVQVQTRAILHIPRIROGLERALL 863

RESULT 13
Q7SVL5
ID Q7SVL5 PRELIMINARY; PRT; 852 AA.
AC Q7SVL5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1];
RP SEQUENCE FROM N.A.
RX DOI=10.1128/JVI.77.11.6359-6366.2003;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
  Brown T.W., Salemi M., Vandamme A.M., Kalish M.L.;
  "U.S. Human immunodeficiency virus type 1 epidemic: date of origin,
  RT population history, and characterization of early strains.";
  J. Virol. 77:6359-6366(2003).
DR EMBL; AY247220; AAP37148.1; -.
DR HSP; P04578; IDLB.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
  Transmembrane.
SQ SEQUENCE 852 AA; 96680 MW; F8D811028025D998 CRC64;

Query Match 84.2%; Score 3886; DB 2; Length 852;
Best Local Similarity 84.6%; Pred. No. 4.5e-279;
Matches 733; Conservative 50; Mismatches 69; Indels 14; Gaps 5;

QY 1 MRVKGIRRNQHWGWTMLGLMLTCSATEKLWTVTVYGVVPVWKEATTTLPFCASDAKAY 60
DB 1 MRVKEIGKNTQYUWRGWTMLGLMLTCSAEQULWTVTVYGVVPVWKEATTTLPFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQEVVELNVVTENFNWKNMVEQWHEDIISLWDQSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQEVGLENVTFENFNWKNMVEQWHEDIISLWDQSLKPCVK 120
QY 121 LTPCLVTNLCTDLGNTTNT-----TSSGGGMEGEIKNCSFNITTSIRDVKQKEYAL 173
```

```
DB 121 LTPCLVTNLCTDLGNTTNT-----TSSGGGMEGEIKNCSFNITTSIRDVKQKEYAL 173
QY 181 LYKLDIEPIDDNNTSVRLISCNSTSVITQACPKEIPPIHYCAPAGFALLKCNCKKFSG 240
DB 174 LYKLDIVPIDDNNTSVRLISCNSTSVITQACPKEIPPIHYCAPAGFALLKCRDKKFKG 233
QY 241 KGSKCNVSTVQCTHGIRPVVSTQLLNGSLABEEVVRSENFTNNAKTIIVOLREPVKIN 300
DB 234 KGPCNTVSTVQCTHGIRPVVSTQLLNGSLABEEVVRSDNFTDNKTIIVOLKESVEIN 293
QY 301 CSRPNNTKSGIPMGPRAFYTTGIIIGDIRQAHCMISKTNWNLAKQVVEKLGGEFKNKT 360
DB 294 CTRPNNTKSHIGPGRAFYTGDIVGDIRQAHCMISKAKWNTTLKQIVTKLREQFGNR 353
QY 361 KIVFTNSSGDDPEIVTHSFNCAGEFFYCNLTQFDSIWNSENGTWNITGLNNTGRNDTI 420
DB 354 TIVFNOSGDDPEIVHMSFCGGEFFYCNLTQFNSWNA-SSTWMDTSGSNNT--EGTI 410
QY 421 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
DB 411 TLPCKRIKQIINRWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGSNEND--TEIFRPG 467
QY 481 GGDMDNRWSELYKYKVKVIEPLGVAPTAKRVRVQREKRAVGIGAVFLGFLGAAGSTWG 540
DB 468 GGNMDNRWSELYKYKVKVIEPLGVAPTAKRVRVQREKRAVGIGAVFLGFLGAAGSTWG 527
QY 541 AASVTLTVQAROLLGSIIVQOQNLLRAEAQOHLQLTVMGKIQARILAVERYLKDOQ 600
DB 528 AASMTLTVQAROLLGSIIVQOQNLLRAEAQOHLQLTVMGKIQARILAVERYLKDOQ 587
QY 601 LLAGWCSGKLICTTTPVWNASWKNKTLKLEAIWNNWMTWQWDKEIDNTSLIYSLIESP 660
DB 588 LLAGWCSGKLICTTTPVWNASWS-NKSLDNIWNNWMTWQWDKEIDNTSLIYSLIESP 646
QY 661 IQOEKNEBELLELDKWNILWNNWFDISNWLWYIKIFIMIVGGLVGLRIVFVLSIVNRVQ 720
DB 647 NQOEKNEBELLELDKWNILWNNWFDITKWLWYIKIFIMIVGGLVGLRIVFVLSIVNRVQ 706
QY 721 GYSPISLQTRLPAPGPRDPEIEEGDRDRDRSGLLVGFLTLIWDLSRLCLFSYHR 780
DB 707 GYSPISLQTRLPAPGPRDPEIEEGDRDRDRSGLLVGFLTLIWDLSRLCLFSYHR 766
QY 781 LRDLILLIYTRIVELLGRGWELIKYWNLLQYWSQELKNSAVLSFNATAIAVAEGTDRVI 840
DB 767 LRDLILLIYTRIVELLGRGWELIKYWNLLQYWSQELKNSAVLSFNATAIAVAEGTDRVI 826
QY 841 QVLQVRGRALLHIPRIROGLERALL 866
DB 827 EVLQRAYRALLHIPRIROGLERALL 852

RESULT 14
O40222 PRELIMINARY; PRT; 854 AA.
AC O40222;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polyprotein.
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1];
RP SEQUENCE FROM N.A.
RX STRAIN=AD8;
RA Theodore T.S., England G., Buckler-White A., Buckler C.E.,
  Martin M.A., Peden K.W.;
  "Construction and characterization of a stable full-length macrophage-
  tropic HIV type 1 molecular clone that directs the production of high
  titers of progeny virions.";
```

```
RL AIDS Res. Hum. Retroviruses 12:191-194 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.C.;
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR AF004394; AAB64170.1; -.
DR HSSP; P04578; IDLB.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR011010; DNA_brk_join_enz.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97291 MW; 06C45E69103C6C12 CRC64;

Query Match      84.2%; Score 3886; DB 2; Length 854;
Best Local Similarity 84.3%; Pred. No. 4.5e-279;
Matches 732; Conservative 47; Mismatches 73; Indels 16; Gaps 8;

QY 1 MRVKGIRRYQHWWGCTMLLGLLMICSAATEKLVVTVYGVVWKEATTLFCASDAKAY 60
DB 1 MKVKGIRKRYQHLLWKGIMLLGLMLCSAVENLVTVYGVVWKEATTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQEVLEVNVTENFNWKNMVEQMHEDIISLDQSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQEVLEVNVTENFNWKNMVEQMHEDIISLDQSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATSTIGDRMKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNNNS-----SEG--MRGEIKNCNFNATSTIGDRMKEYAL 171
QY 181 LYKLDIEPIDNDNTSVRLISCVNTSVITQACPKISFEPPIHYCAPAGFAILKCNKKFSG 240
DB 172 FYRLDVVPIDNDNTSVRLINCNTSTITQACPKVSFEPPIHYCTPAGFAILKCKKKFNG 231
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLREPVKIN 300
DB 232 TGPCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLKESVEIN 291
QY 301 CSRPNNTNRKSIIPMGPRAFYTTGQIIGDIIROAHCNISKNTNNAKTIIVQLREPVKIN 360
DB 292 CTRPNNTNRKSIHIGPRAFYTTGQIIGDIIROAHCNISRKNNTNNAKTIIVQLKESVEIN 351
QY 361 K-IVFTNSSGGDPEIVTHSFNCAGEFFYCNNTQTFDSIWNSENGTWNITRGLNNTGRNDT 419
DB 352 KTVFVNSGGDPEIVHNSFNCAGEFFYCNNTQTFDSIWNSENGTWNITRGLNNTGRNDT 409
QY 420 ITLPCRKQIINRWQEVGKAYAPPIKGNISCSNNTIGLLTRDGGKDDNSRDGNETFRP 479
DB 410 ITLPCRKQIINRWQEVGKAYAPPIKGNISCSNNTIGLLTRDGG--NNHNDTETFRP 467
QY 480 GGGDMRDNRSELYKYKVKVIEPLGVAPTAKRVRVQREERAVG-LGAMPFGFLGAAGST 538
DB 468 GGGDMRDNRSELYKYKVKVIEPLGVAPTAKRVRVQREERAVGTTGAMPLGFLGAAGST 527
QY 539 MGAASVTLTVQARQLSGIVQOQSNLLRAEAQOHLQLTWGIGKQIARILAVERYLKD 598
DB 528 MGAASVTLTVQARQLSGIVQOQSNLLRAEAQOHLQLTWGIGKQIARILAVERYLKD 587
QY 599 QQLLGLWGCSGKLCITTVPPNNAWSKNKTLKALNNMTWMDKEIDNVTSLIYSIER 658
DB 588 QQLLGLWGCSGKLCITTVPPNNAWS--NKTLDIWNMTWMDKEIDNVTSLIYSIER 646
QY 659 SPIQEQNEQELLELDKAWLWFWFDSIWNMTWIKIPIMVGLVGLIRIVFVLSIVNRV 718
DB 647 SQNQEQNEQELLELDKAWLWFWFDSIWNMTWIKIPIMVGLVGLIRIVFVLSIVNRV 706
QY 719 RQGYSPLSFOTRLPAPRGPDRPEIEEGGDRDRSGLLVDFGLTLIIVDRLSLCLFSY 778
DB 707 RQGYSPLSFOTRLPAPRGPDRPEIEEGGDRDRSGLLVDFGLTLIIVDRLSLCLFSY 766
QY 779 HRLRDLILLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 838
DB 767 HRLRDLILLIVARIVELLGRRGWEALKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 826
QY 839 VIQVLRQVGRALLHIPTIRIQGLERALL 866
DB 827 VIEIVQIRAILHIPTIRIQGLERALL 854

RESULT 15
Q03811 ID Q03811 PRELIMINARY; PRT; 853 AA.
AC Q03811;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface envelope glycoprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91195299; PubMed=2014229;
RA Westervelt P., Gendelman H.E., Ratner L.;
RT "Identification of a determinant within the human immunodeficiency
RT virus 1 surface envelope glycoprotein critical for productive
RT infection of primary monocytes";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101(1991).
DR EMBL; M60472; AAA45065.1; -.
DR HSSP; P20871; 1CE4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 853 AA; 96998 MW; AD2AF21E2B06AD78 CRC64;

Query Match      83.8%; Score 3870.5; DB 2; Length 853;
Best Local Similarity 83.9%; Pred. No. 6.4e-278;
Matches 728; Conservative 52; Mismatches 71; Indels 17; Gaps 8;

QY 1 MRVKGIRRYQHWWGCTMLLGLLMICSAATEKLVVTVYGVVWKEATTLFCASDAKAY 60
DB 1 MKVKGIRKRYQHLLWKGIMLLGLMLCSAVENLVTVYGVVWKEATTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQEVLEVNVTENFNWKNMVEQMHEDIISLDQSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQEVLEVNVTENFNWKNMVEQMHEDIISLDQSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATSTIGDRMKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNNNS-----SEG--MRGEIKNCNFNATSTIGDRMKEYAL 171
QY 181 LYKLDIEPIDNDNTSVRLISCVNTSVITQACPKISFEPPIHYCAPAGFAILKCNKKFSG 240
DB 172 FYRLDVVPIDNDNTSVRLINCNTSTITQACPKVSFEPPIHYCTPAGFAILKCKKKFNG 231
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLREPVKIN 300
DB 232 TGPCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLKESVEIN 291
QY 301 CSRPNNTNRKSIIPMGPRAFYTTGQIIGDIIROAHCNISKNTNNAKTIIVQLREPVKIN 360
DB 292 CTRPNNTNRKSIHIGPRAFYTTGQIIGDIIROAHCNISRKNNTNNAKTIIVQLKESVEIN 351
QY 361 K-IVFTNSSGGDPEIVTHSFNCAGEFFYCNNTQTFDSIWNSENGTWNITRGLNNTGRNDT 419
```

Db	352	KTIVNQSGGDEIIVHSENCGEFFYCNSITQFNSTWNF-NGTWNLTQS-NGTEGNDT	409
Qy	420	ITLPCRICKIINWQEVGKAMYAPPIKGNISCSSNITGLLLTRDCKDDNSRDGNETFRP	479
Db	410	ITLPCRICKIINWQEVGKAMYAPPIRGQIRCSSNITGLILTRDGGTNS--GSEIFRP	466
Qy	480	GGGDMRDNRSELYKYKVVKIEPLGVAPTAKKRRVVQREERAVG-LGAMFFGFLGAAGST	538
Db	467	GGGDMRDNRSELYKYKVVKIEPLGVAPTAKKRRVVQREKRAVGTIGAMFLGFLGAAGST	526
Qy	539	MGAASVTLTVQAROLLSGIVQQSNLLRAIEAQCHLLQLTWNGIKOLQARILAVERYLKD	598
Db	527	MGAASITLTQARLLLSGIVQQSNLLRAIEAQCHLLQLTWNGIKOLQARVLALERYLKD	586
Qy	599	QQLLGWGCCKLICITTVPNWNASWKNKLTLEAIWNNMTWMDWDEIDNYTSLIYSLIEE	658
Db	587	QQLLGWGCCKLICITTVPNWNASW-NKTLDMWDNMTWMEWEREIENTYGLIYTLIEE	645
Qy	659	SPIQKNEQELLELDKWNLMNWFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRV	718
Db	646	SONQKNEQDILLALDKWASLWNFDISNWLWYIKIFIMIVGGLIGLRIVFTVLSVNRV	705
Qy	719	RQGYSPLSFOTRPLPAPRGDRPREEEGGDRDRDRSGLLVGDFLTIWVDLRSCLFSY	778
Db	706	RQGYSPLSFOTRPLPAPRGDRPREEEGGDRDRDRSGLLVGDFLTIWVDLRSCLFSY	765
Qy	779	HLRDLILLIVTRIVELLGREGWELKYWNLLQYWSQELKNSAVSLFNATAVAEGTDR	838
Db	766	HLRDLILLIVARIVELLGRRGEVLKYWNLLQYWSQELRNSAVSLFNATAVAEGTDR	825
Qy	839	VIOVQVRGALLHIPTIRIQGLERALL	866
Db	826	VIEVQRIYRAILHIPTIRIQGLERULL	853

Search completed: April 11, 2005, 11:14:03  
Job time : 282.268 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:37:31 ; Search time 11.4674 Seconds  
(without alignments)  
1180.438 Million cell updates/sec

Title: US-09-762-261C-3  
Perfect score: 195  
Sequence: 1 CSRPNNTKSPMGPRAPFTTGTQIIIGDIOAHC 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	195	100.0	35	3 AAY79023	Aay79023 HIV-1 RV
2	195	100.0	866	3 AAY79020	Aay79020 HIV-1 env
3	185	94.9	35	2 AAR20883	Aar20883 PND EE378
4	185	94.9	35	2 AAR20953	Aar20953 PND EE669
5	185	94.9	35	2 AAR20853	Aar20853 PND EE322
6	183	93.8	35	2 AAR21134	Aar21134 PND EE228
7	183	93.8	35	2 AAR20911	Aar20911 PND EE533
8	183	93.8	35	2 AAR21135	Aar21135 PND EE228
9	183	93.8	35	2 AAR20912	Aar20912 PND EE533
10	183	93.8	35	2 AAR20910	Aar20910 PND EE533
11	183	93.8	35	2 AAR21136	Aar21136 PND EE228
12	183	93.8	35	2 AAR29113	Aar29113 Group II
13	181	92.8	842	3 AAB69350	Aab69350 HIV-1 non
14	180	92.3	35	2 AAR20945	Aar20945 PND EE665
15	180	92.3	35	2 AAR20947	Aar20947 PND EE665
16	180	92.3	35	2 AAR20946	Aar20946 PND EE665
17	180	92.3	35	2 AAR29110	Aar29110 Group II
18	180	92.3	35	2 AAR29117	Aar29117 Group II
19	179	91.8	35	3 AAB10703	Aab10703 HIV-1 iso
20	179	91.8	35	5 ABB73683	Abb73683 HIV-1 V3-
21	178	91.3	35	2 AAR20936	Aar20936 PND EE639
22	177	90.8	35	8 ADR50777	Adr50777 Antigenic
23	177	90.8	45	8 ADR58150	Adr58150 Novel ant
24	176	90.3	35	2 AAR20881	Aar20881 PND EE378
25	176	90.3	35	2 AAR21092	Aar21092 PND EE377

26	176	90.3	35	2 AAR20837	Aar20837 PND EE304
27	176	90.3	35	2 AAR20840	Aar20840 PND EE308
28	176	90.3	35	2 AAR20851	Aar20851 PND EE320
29	176	90.3	35	2 AAR20882	Aar20882 PND EE378
30	176	90.3	35	2 AAR20941	Aar20941 PND EE661
31	176	90.3	35	2 AAR68030	Aar68030 HIV-1 sub
32	176	90.3	35	2 AAR66751	Aar66751 HIV-1 gpl
33	176	90.3	35	2 AAW07390	Aaw07390 HIV-1 gpl
34	176	90.3	35	2 AAW03356	Aaw03356 Human imm
35	176	90.3	35	2 AAW39914	Aaw39914 V3 sequen
36	176	90.3	35	2 AAY31161	Aay31161 HIV-1 pep
37	176	90.3	35	3 AAB38961	Aab38961 HIV-1 pep
38	176	90.3	35	4 AAB71933	Aab71933 HIV-1 gpl
39	176	90.3	35	4 AAB66184	Aab66184 HIV-1 V3
40	176	90.3	35	8 ADI26505	Adi26505 Peptide o
41	176	90.3	35	8 ADR50753	Adr50753 Antigenic
42	176	90.3	37	2 AAR41065	Aar41065 HIV-1 com
43	176	90.3	39	2 AAW84056	Aaw84056 HIV-1 Ba-
44	176	90.3	43	2 AAW63735	Aaw63735 HIV-1 JR-
45	176	90.3	45	8 ADR58145	Adr58145 Novel ant

## ALIGNMENTS

RESULT 1  
AAY79023  
ID AAY79023 standard; protein; 35 AA.  
XX  
AC AAY79023;  
XX  
DT 12-SEP-2003 (revised)  
DT 05-JUN-2000 (first entry)  
XX  
DE HIV-1 RV strain envelope protein V3 region peptide sequence #2.  
XX  
KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;  
KW vaccine; antibody; V3 region; detect.  
XX  
OS Human immunodeficiency virus 1.  
XX  
PN WO200007631-A1.  
XX  
PD 17-FEB-2000.  
XX  
PF 04-AUG-1999; 99WO-US017596.  
XX  
PR 04-AUG-1998; 98US-0095267P.  
XX  
(JACK-) JACKSON FOUND HENRY M.  
Quinnan GV, Zhang PF;  
WPI; 2000-205578/18.  
XX  
Human immunodeficiency virus-1 envelope protein, useful as a vaccine or immunogenic preparation against HIV-1 infection induces the production of broadly cross-reactive neutralizing anti-serum.  
XX  
Example 5; Page 27; 54pp; English.  
XX  
This sequence represents a human immunodeficiency virus type 1 (HIV-1) R2 strain envelope (env) protein V3 region peptide sequence. The invention relates to the HIV-1 env protein or its fragments, which when administered to a mammal, induces the production of broadly cross-reactive neutralising anti-serum against multiple strains of HIV-1. The HIV-1 env protein and its fragments are useful as vaccines, immunogenic compositions or diagnostic reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing antibodies in a sample. The amino acid sequence and its fragments or mRNA sequence are useful for generating antibodies in a mammal. In addition, a recombinant delivery vector containing the env amino acid sequence may also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS field)

```

XX SQ Sequence 35 AA;
Query Match 100.0%; Score 195; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35
| | | | | | | | | | | | | | | | | | | | |
Db 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35

RESULT 2
AA79020
ID AAY79020 standard; protein; 866 AA.
XX AC AAY79020;
XX DT 12-SEP-2003 (revised)
XX DT 05-JUN-2000 (first entry)
XX DE HIV-1 envelope protein amino acid sequence.
XX KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;
XX KW vaccine; antibody; detect.
XX OS Human immunodeficiency virus 1.
XX PN W0200007631-A1.
XX PD 17-FEB-2000.
XX PF 04-AUG-1999; 99WO-US017596.
XX PR 04-AUG-1998; 98US-0095267P.
XX PA (JACK-) JACKSON FOUND HENRY M.
XX PI Quinnan GV, Zhang PF;
XX DR WPI; 2000-205578/18.
XX PT Human immunodeficiency virus-1 envelope protein, useful as a vaccine or
PT immunogenic preparation against HIV-1 infection induces the production of
PT broadly cross-reactive neutralizing anti-serum.
XX PS Claim 2; Page 33; 54pp; English.
XX CC This sequence represents a human immunodeficiency virus type 1 (HIV-1)
CC envelope (env) protein amino acid sequence. The invention relates to the
CC HIV-1 env protein or its fragments, which when administered to a mammal,
CC induces the production of broadly cross-reactive neutralising anti-serum
CC against multiple strains of HIV-1. The HIV-1 env protein and its
CC fragments are useful as vaccines, immunogenic compositions or diagnostic
CC reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing
CC antibodies in a sample. The amino acid sequence and its fragments or mRNA
CC sequence are useful for generating antibodies in a mammal. In addition, a
CC recombinant delivery vector containing the env amino acid sequence may
CC also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS
XX field)
XX SQ Sequence 866 AA;
Query Match 100.0%; Score 195; DB 3; Length 866;
Best Local Similarity 100.0%; Pred. No. 3.1e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35
| | | | | | | | | | | | | | | | | | | | |
Db 301 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 335

RESULT 3

```

```

AAR20883
XX ID AAR20883 standard; protein; 35 AA.
XX AC AAR20883;
XX DT 01-JUN-1992 (first entry)
XX DE PND EE378-3.
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX OS Human immunodeficiency virus.
XX PN EP471407-A.
XX PD 19-FEB-1992.
XX PF 07-AUG-1991; 91EP-00202025.
XX PR 13-AUG-1990; 90US-00566638.
XX PR 13-AUG-1990; 90US-00566654.
XX PR 13-AUG-1990; 90US-00566656.
XX PA (MERI ) MERCK & CO INC.
XX PI Lewis JA, Davide JP, Waterbury JA;
XX DR WPI; 1992-058471/08.
XX DR N-PSDB; AAQ21298.
XX PT New antigenic conjugate of HIV major neutralisation determinant -
XX PT covalently linked to outer membrane proteosome of Neisseria, useful as
XX PT vaccine and in treating and preventing HIV infections, AIDS and ARC.
XX PS Claim 1; Page 175 + 117-118; 177pp; English.
XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
XX CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
XX CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
XX CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,
XX CC then linked by cross-linking agents, monogeneric spacers or bigeneric
XX CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
XX CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
XX CC Glu or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
XX CC post- exposure to prevent or treat HIV infection or disease, and are
XX CC capable of producing an immune response specific for the immunogen. See
XX CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78
XX SQ Sequence 35 AA;
Query Match 94.9%; Score 185; DB 2; Length 35;
Best Local Similarity 91.4%; Pred. No. 2.5e-17;
Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35
| | | | | | | | | | | | | | | | | | | | |
Db 1 CTRPNNTKRSIPGPGAFYTTGQIIGDIRQAHC 35

RESULT 4
AAR20953
XX ID AAR20953 standard; protein; 35 AA.
XX AC AAR20953;
XX DT 01-JUN-1992 (first entry)
XX DE PND EE669-3.
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX SQ

```



OS Human immunodeficiency virus.  
 XX EP471407-A.  
 XX  
 PD 19-FEB-1992.  
 XX  
 XX 07-AUG-1991; 91EP-00202025.  
 PF  
 XX 13-AUG-1990; 90US-00566638.  
 PR 13-AUG-1990; 90US-00566654.  
 PR 13-AUG-1990; 90US-00566656.  
 XX  
 XX (MERI ) MERCK & CO INC.  
 PA  
 XX Lewis JA, Davide JP, Waterbury JA;  
 XX  
 XX WPI; 1992-058471/08.  
 DR N-PSDB; AAQ21370.  
 DR  
 XX New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.  
 XX  
 XX Claim 1; Page 175 + 168; 177pp; English.  
 PS  
 XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78  
 XX  
 XX SQ Sequence 35 AA;  
 Query Match 94.9%; Score 185; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 2.5e-17;  
 Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35  
 Db :|||||:|||||:|||||:|||||:|||||  
 1 CTRPNNTKRSIPGPGAFYTTGTEIGDIRQAHC 35  
 RESULT 5  
 AAR20853  
 ID AAR20853 standard; protein; 35 AA.  
 XX  
 AC AAR20853;  
 XX  
 DT 01-JUN-1992 (first entry)  
 XX  
 DE PND EE322-2.  
 XX  
 KW HIV; PND; principal neutralising determinant; Omp; Neisseria;  
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.  
 XX  
 OS Human immunodeficiency virus.  
 XX  
 XX EP471407-A.  
 PN  
 XX 19-FEB-1992.  
 PD  
 XX 07-AUG-1991; 91EP-00202025.  
 PF  
 XX 13-AUG-1990; 90US-00566638.  
 PR 13-AUG-1990; 90US-00566654.  
 PR 13-AUG-1990; 90US-00566656.  
 XX  
 XX

PA (MERI ) MERCK & CO INC.  
 XX Lewis JA, Davide JP, Waterbury JA;  
 XX  
 XX WPI; 1992-058471/08.  
 DR N-PSDB; AAQ21268.  
 DR  
 XX New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.  
 XX  
 XX Claim 1; Page 175 + 93-94; 177pp; English.  
 PS  
 XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78  
 XX  
 XX SQ Sequence 35 AA;  
 Query Match 94.9%; Score 185; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 2.5e-17;  
 Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35  
 Db :|||||:|||||:|||||:|||||:|||||  
 1 CTRPNNTKRSIPGPGAFYTTGTEIGDIRQAHC 35  
 RESULT 6  
 AAR21134  
 ID AAR21134 standard; protein; 35 AA.  
 XX  
 AC AAR21134;  
 XX  
 DT 01-JUN-1992 (first entry)  
 XX  
 DE PND EE228-1.  
 XX  
 KW HIV; PND; principal neutralising determinant; Omp; Neisseria;  
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.  
 XX  
 OS Human immunodeficiency virus.  
 XX  
 XX EP471407-A.  
 PN  
 XX 19-FEB-1992.  
 PD  
 XX 07-AUG-1991; 91EP-00202025.  
 PF  
 XX 13-AUG-1990; 90US-00566638.  
 PR 13-AUG-1990; 90US-00566654.  
 PR 13-AUG-1990; 90US-00566656.  
 XX  
 XX (MERI ) MERCK & CO INC.  
 PA  
 XX Lewis JA, Davide JP, Waterbury JA;  
 XX  
 XX WPI; 1992-058471/08.  
 DR N-PSDB; AAQ21232.  
 DR  
 XX New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.  
 XX  
 XX Claim 1; Page 175 + 64; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;  
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 |:|||||:|||||:|||||:|||||:|||||  
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35

RESULT 7

AAQ20911  
 ID AAR20911 standard; protein; 35 AA.

XX AC AAR20911;

XX DT 01-JUN-1992 (first entry)

XX DE PND EE533-2.

XX HW; PND; principal neutralising determinant; Omp; Neisseria;  
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX PN EP471407-A.

XX PD 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX PA (MERI ) MERCK & CO INC.

XX PI Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21335.

XX PT New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX PS Claim 1; Page 175 + 140; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;  
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 |:|||||:|||||:|||||:|||||:|||||  
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35

RESULT 8

AAQ21135  
 ID AAR21135 standard; protein; 35 AA.

XX AC AAR21135;

XX DT 01-JUN-1992 (first entry)

XX DE PND EE228-2.

XX HW; PND; principal neutralising determinant; Omp; Neisseria;  
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX PN EP471407-A.

XX PD 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX PA (MERI ) MERCK & CO INC.

XX PI Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21233.

XX PT New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX PS Claim 1; Page 175 + 65; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;  
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 |:|||||:|||||:|||||:|||||:|||||  
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35

RESULT 9  
AAR20912  
ID AAR20912 standard; protein; 35 AA.

XX AC AAR20912;  
XX DT 01-JUN-1992 (first entry)  
XX DE PND EE533-3.  
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;  
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.  
XX OS Human immunodeficiency virus.

XX EP471407-A.  
XX PN 19-FEB-1992.  
XX PD  
XX PF 07-AUG-1991; 91EP-00202025.  
XX PR 13-AUG-1990; 90US-00566638.  
XX PR 13-AUG-1990; 90US-00566654.  
XX PR 13-AUG-1990; 90US-00566656.  
XX PA (MERI ) MERCK & CO INC.  
XX PI Lewis JA, Davide JP, Waterbury JA;  
XX WPI; 1992-058471/08.  
XX DR N-PSDB; AAQ21336.

XX New antigenic conjugate of HIV major neutralisation determinant -  
PT covalently linked to outer membrane proteosome of Neisseria, useful as  
PT vaccine and in treating and preventing HIV infections, AIDS and ARC.  
XX Claim 1; Page 175 + 141; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,  
CC then linked by cross-linking agents, monogenic spacers or bigeneric  
CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
CC Glu or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
CC post- exposure to prevent or treat HIV infection or disease, and are  
CC capable of producing an immune response specific for the immunogen. See  
CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;  
Best Local Similarity 91.4%; Pred. No. 4.6e-17;  
Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHC 35  
|:|||||||||:|||||||||:|||||||||:|||||||||  
Db 1 CTRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHC 35

RESULT 10

AAR20910  
ID AAR20910 standard; protein; 35 AA.

XX AC AAR20910;  
XX DT 01-JUN-1992 (first entry)  
XX DE PND EE533-1.

XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;  
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX EP471407-A.

XX PN 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX PA (MERI ) MERCK & CO INC.

XX PI Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX DR N-PSDB; AAQ21334.

XX New antigenic conjugate of HIV major neutralisation determinant -  
PT covalently linked to outer membrane proteosome of Neisseria, useful as  
PT vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 139-140; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,  
CC then linked by cross-linking agents, monogenic spacers or bigeneric  
CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
CC Glu or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
CC post- exposure to prevent or treat HIV infection or disease, and are  
CC capable of producing an immune response specific for the immunogen. See  
CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;  
Best Local Similarity 91.4%; Pred. No. 4.6e-17;  
Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHC 35  
|:|||||||||:|||||||||:|||||||||:|||||||||

Db 1 CTRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHC 35

RESULT 11

AAR21136  
ID AAR21136 standard; protein; 35 AA.

XX AC AAR21136;

XX DT 01-JUN-1992 (first entry)

XX DE PND EE238-3.

XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;  
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX EP471407-A.

XX PN 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX (MERI ) MERCK & CO INC.  
 PA Lewis JA., Davide JP, Waterbury JA;  
 PI WPI; 1992-058471/08.  
 PT N-PSDB; AAQ21234.  
 DR  
 XX  
 XX New antigenic conjugate of HIV major neutralisation determinant -  
 PT covalently linked to outer membrane proteosome of Neisseria, useful as  
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.  
 XX  
 PS Claim 1; Page 175 + 66; 177pp; English.  
 XX  
 CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78  
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for  
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic  
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,  
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric  
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids  
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,  
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or  
 CC post- exposure to prevent or treat HIV infection or disease, and are  
 CC capable of producing an immune response specific for the immunogen. See.  
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78  
 XX  
 XX Sequence 35 AA;  
 SQ  
 Query Match 93.8%; Score 183; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;  
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 RESULT 12  
 AAR29113  
 ID AAR29113 standard; protein; 35 AA.  
 XX  
 AC AAR29113;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 02-APR-1993 (first entry)  
 XX  
 DE Group II HIV gp120 PND peptide 8926C.  
 XX  
 KW Principal Neutralising determinant; Human immunodeficiency virus;  
 KW vaccine; Robson's analytical method; polymerase chain reaction;  
 KW Garnier-Osguthorpe-Robson method; GOR method; secondary structure.  
 XX  
 OS Human immunodeficiency virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 8..14  
 FT /note= "classification of PND peptides is determined by  
 FT the predicted secondary structure of this region - see  
 FT comments"  
 FT 15..18  
 FT /note= "conserved PND motif"  
 FT  
 XX  
 XX EP516135-A2.  
 XX  
 PD 02-DEC-1992.  
 XX  
 XX 29-MAY-1992; 92BP-00109072.  
 XX  
 XX 31-MAY-1991; 91JP-00129224.  
 XX  
 XX (KAGA ) CHENO SERO THERAPEUTIC RES INS.  
 PA  
 PI Eda Y, Shiosaki K, Osatomi K, Tokiyoshi S;  
 XX

XX WPI; 1992-400517/49.  
 XX  
 PT Principle neutralising determinant peptide(s) of HIV gp120 protein - used  
 PT for diagnosing, preventing and treating HIV infection.  
 XX  
 XX Example 1; Page 10; 26pp; English.  
 PS  
 XX DNA encoding HIV PND peptides was PCR amplified using genomic DNA from  
 CC HIV-infected peripheral blood mononuclear cells as template. The  
 CC amplified fragments were fused to beta-galactosidase coding sequence.  
 CC E.coli transformants were cultured to produce the fusion protein. The  
 CC expressed PND proteins were divided into groups based on their reactivity  
 CC with neutralising antibodies and their amino acid sequence. The amino  
 CC acid sequence was analysed using Robson's analytical program for protein  
 CC secondary structure. Five groups were identified and 90% of all  
 CC previously reported PND peptides were included in 3 main groups (i.e.  
 CC Groups I, II and III). Group II PND peptides are those which have the  
 CC structure XXBBBX on the amino-terminal side of the GPGR motif (B = beta-  
 CC strand structure and X = turn or coil structure). Vaccine preparations  
 CC comprising representative peptides from each of the 5 groups can be used  
 CC to develop vaccines able to recognise any HIV variant. See AAQ31607-  
 CC Q31608, AAR28995-R29000 and AAR29110-R29128. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX  
 XX Sequence 35 AA;  
 SQ  
 Query Match 93.8%; Score 183; DB 2; Length 35;  
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;  
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35  
 RESULT 13  
 AAB69350  
 ID AAB69350 standard; protein; 842 AA.  
 XX  
 AC AAB69350;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 20-APR-2001 (first entry)  
 XX  
 DE HIV-1 non-subtype B clone 93BR029-4 env protein.  
 XX  
 KW HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env; vpu;  
 KW vif; vpr; tat; rev; nef; vaccine.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO2000026416-A1.  
 XX  
 PD 11-MAY-2000.  
 XX  
 PF 25-OCT-1999; 99WO-US024837.  
 XX  
 PR 02-NOV-1998; 98US-00184418.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 XX Hahn BH, Shaw GM, Gao F;  
 XX  
 XX WPI; 2000-365651/31.  
 XX  
 XX Novel genomic nucleic acids of non-subtype B human immunodeficiency virus  
 PT type 1 useful for detecting and treating AIDS comprises a specific  
 PT nucleotide sequence.  
 XX  
 XX Claim 41; Fig 21; 131pp; English.  
 PS  
 XX The present in invention provides the protein and coding sequences for a

Query Match 92.3%; Score 180; DB 2; Length 35;  
Best Local Similarity 88.6%; Pred. No. 1.2e-16;

Search completed: April 11, 2005, 11:09:03  
Job time : 12.4674 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:57:57 ; Search time 2.65208 Seconds  
(without alignments)  
1269.790 Million cell updates/sec

Title: US-09-762-261C-3

Perfect score: 195

Sequence: 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	92.3	35	2	PC2296
2	176	90.3	35	2	PC2295
3	176	90.3	77	2	S35862
4	176	90.3	110	2	A46410
5	176	90.3	847	2	T09448
6	176	90.3	847	2	S13289
7	175	89.7	76	2	S35859
8	175	89.7	77	2	S35820
9	175	89.7	77	2	S35819
10	174	89.2	443	2	C41621
11	173	88.7	77	2	S35834
12	173	88.7	77	2	S35836
13	173	88.7	77	2	S35835
14	172	88.2	77	2	S35848
15	172	88.2	77	2	S35849
16	172	88.2	77	2	S35857
17	171	87.7	35	2	PC2297
18	171	87.7	46	2	G01847
19	171	87.7	77	2	S35823
20	171	87.7	77	2	S35809
21	171	87.7	110	2	B46410
22	171	87.7	110	2	C46410
23	171	87.7	290	2	S25940
24	171	87.7	445	2	A41621
25	170	87.2	77	2	S35806
26	170	87.2	77	2	S35841
27	170	87.2	77	2	S35816
28	170	87.2	77	2	S35804
29	170	87.2	77	2	S35850

30	170	87.2	843	1	H44001	env polyprotein pr
31	169	86.7	71	2	S42918	env polyprotein -
32	169	86.7	77	2	S35821	envelope protein -
33	168	86.2	77	2	S35824	envelope protein -
34	168	86.2	77	2	S35825	envelope protein -
35	168	86.2	110	2	D46410	envelope protein (
36	168	86.2	852	1	VCLJBR	env polyprotein -
37	167	85.6	71	2	S42904	env polyprotein -
38	167	85.6	71	2	S42905	env polyprotein -
39	167	85.6	71	2	S42898	env polyprotein -
40	167	85.6	77	2	S35815	envelope protein -
41	166	85.1	506	2	A40218	envelope glycoprote
42	165	84.6	77	2	S35800	envelope protein -
43	165	84.6	77	2	S35801	envelope protein -
44	165	84.6	299	2	S60529	envelope polyprote
45	165	84.6	855	1	VCLJAJ	env polyprotein pr

ALIGNMENTS

RESULT 1

PC2296

V3 domain peptide P4346 - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 21-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 09-Jul-2004

C:Accession: PC2296

R:Sherefa, K.; Soennerborg, A.; Steinbergs, J.; Saellberg, M.

Biochem. Biophys. Res. Commun. 205, 1658-1664, 1994

A:Title: Rapid grouping of HIV-1 infection in subtypes A to E by V3 peptide serotyping ar

A:Reference number: PC2291; MUID:95110306; PMID:7811250

A:Accession: PC2296

A:Molecule type: protein

A:Residues: 1-35 <SHE>

A:Cross-references: UNIPROT:Q70826; UNIPROT:Q70831

C:Superfamily: Type E retrovirus env polyprotein

Query Match 92.3%; Score 180; DB 2; Length 35;

Best Local Similarity 91.4%; Pred. No. 2.2e-18;

Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35

DB 1 CTRPNNTKRSIHMGPGRAFYTTGTEIGDIRQAHC 35

RESULT 2

PC2295

V3 domain peptide P4317 - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 21-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 09-Jul-2004

C:Accession: PC2295

R:Sherefa, K.; Soennerborg, A.; Steinbergs, J.; Saellberg, M.

Biochem. Biophys. Res. Commun. 205, 1658-1664, 1994

A:Title: Rapid grouping of HIV-1 infection in subtypes A to E by V3 peptide serotyping ar

A:Reference number: PC2291; MUID:95110306; PMID:7811250

A:Accession: PC2295

A:Molecule type: protein

A:Residues: 1-35 <SHE>

A:Cross-references: UNIPROT:Q8JA81; UNIPROT:Q8JA72; UNIPROT:Q8JA67; UNIPROT:Q8JA75; UNIPR

C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 35;

Best Local Similarity 88.6%; Pred. No. 8e-18;

Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35

DB 1 CTRPNNTKRSIHMGPGRAFYTTGTEIGDIRQAHC 35

RESULT 3

S35862

envelope protein - human immunodeficiency virus type 1 (strain CSF2951) (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S35862; S35860; S35863  
R:Chiodi, F.  
submitted to the EMBL Data Library, June 1993  
A:Reference number: S35800  
A:Accession: S35862  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-77 <CHI>  
A:Cross-references: UNIPROT:Q77544; UNIPROT:Q77542; UNIPROT:Q77545; EMBL:Z23246; NID:g31362  
A:Experimental source: strain CSF2951 clone3; strain CSF2951 clone4  
A:Accession: S35860  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-77 <CH2>  
A:Cross-references: EMBL:Z23244; NID:g313631; PIDN:CAA80762.1; PID:g313632; EMBL:Z23247;  
A:Experimental source: strain CSF2951 clone11; strain CSF2951 clone4  
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 77;  
Best Local Similarity 88.6%; Pred. No. 1.8e-17;  
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35  
DB 26 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 60

RESULT 4  
A46410  
envelope protein (V3 domain) - human immunodeficiency virus type 1 (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A46410  
R:McNearney, T.; Hornickova, Z.; Markham, R.; Birdwell, A.; Arens, M.; Saah, A.; Ratner, Proc. Natl. Acad. Sci. U.S.A. 89, 10247-10251, 1992  
A:Title: Relationship of human immunodeficiency virus type 1 sequence heterogeneity to a  
A:Reference number: A46410; MUID:93066216; PMID:1438212  
A:Accession: A46410  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-110 <MCN>  
A:Cross-references: UNIPROT:Q9PY06  
A:Experimental source: subject S1  
A>Note: sequence extracted from NCBI backbone (NCBIP:117724)  
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 110;  
Best Local Similarity 88.6%; Pred. No. 2.7e-17;  
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35  
DB 37 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 71

RESULT 5  
T09448  
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: T09448  
R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie, submitted to the EMBL Data Library, July 1996  
A:Reference number: Z16673  
A:Accession: T09448  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-847 <PAN>  
A:Cross-references: UNIPROT:Q75760; EMBL:U63632; NID:g1465777; PID:g1465781  
C:Genetics:

A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 847;  
Best Local Similarity 88.6%; Pred. No. 2.2e-16;  
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35  
DB 293 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 327

RESULT 6  
S13289  
env protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S13289  
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A. Nature 348, 69-73, 1990  
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c  
A:Reference number: S13288; MUID:91043044; PMID:2172833  
A:Accession: S13289  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-847 <OBR>  
A:Cross-references: UNIPROT:Q75760  
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 847;  
Best Local Similarity 88.6%; Pred. No. 2.2e-16;  
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35  
DB 293 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 327

RESULT 7  
S35859  
envelope protein - human immunodeficiency virus type 1 (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S35859; S35871  
R:Chiodi, F.  
submitted to the EMBL Data Library, June 1993  
A:Reference number: S35800  
A:Accession: S35859  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-76 <CHI>  
A:Cross-references: UNIPROT:Q77541; EMBL:Z23243; NID:g313629; PIDN:CAA80761.1; PID:g31363  
C:Superfamily: type E retrovirus env polyprotein

Query Match 89.7%; Score 175; DB 2; Length 76;  
Best Local Similarity 88.6%; Pred. No. 2.5e-17;  
Matches 31; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35  
DB 25 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 59

RESULT 8  
S35820  
envelope protein - human immunodeficiency virus type 1 (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S35820; S35822  
R:Chiodi, F.  
submitted to the EMBL Data Library, June 1993  
A:Reference number: S35800  
A:Accession: S35820



```
A;Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77504; UNIPROT:Q77502; EMBL:Z23201; NID:G313548; PIDN:CAA80
C;Superfamily: type E retrovirus env polyprotein

Query Match      89.7%; Score 175; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 2.5e-17;
Matches 31; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIIGDIRQAH 35
Db 26 CTRPNNTKSIITIGPGRAFYTTGDIIGDIRQAH 60

RESULT 9
S35819
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35819; S35856
R;Chiodi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35819
A;Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77501; UNIPROT:Q77538; EMBL:Z23200; NID:G313546; PIDN:CAA80
C;Superfamily: type E retrovirus env polyprotein

Query Match      89.7%; Score 175; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 2.5e-17;
Matches 31; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIIGDIRQAH 35
Db 26 CTRPNNTKSIITIGPGRAFYTTGDIIGDIRQAH 60

RESULT 10
C41621
env polyprotein P - human immunodeficiency virus type 1 (fragment)
N;Alternate names: coat polyprotein
N;Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 11-Feb-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C41621
R;Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991
A;Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity
A;Reference number: A41621; MUID:92107924; PMID:1763038
A;Accession: C41621
A:Molecule type: DNA
A:Residues: 1-443 <SUR>
A;Cross-references: UNIPROT:Q80023; GB:M77230; NID:G328631; PIDN:AAB03792.1; PID:G555015
A;Note: this virus was isolated from the mother's sexual partner
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro
F;1-251/Product: coat protein gp120 (fragment) #status predicted <GP1>
F;252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>
F;424-443/Domain: transmembrane #status predicted <TM>
F;9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: ca

Query Match      89.2%; Score 174; DB 2; Length 443;
Best Local Similarity 88.6%; Pred. No. 2.2e-16;
Matches 31; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIIGDIRQAH 35
Db 26 CTRPNNTKSIITIGPGRAFYTTGDIIGDIRQAH 60

RESULT 11
S35834
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35834; S35837
R;Chiodi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35834
A;Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77519; UNIPROT:Q77516; EMBL:Z23216; EMBL:Z23219
C;Superfamily: type E retrovirus env polyprotein

Query Match      88.7%; Score 173; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 4.8e-17;
Matches 31; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIIGDIRQAH 35
Db 26 CTRPNNTKSIHMGPGRAFYATGDIIGDIRQAH 60

RESULT 12
S35836
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35836
R;Chiodi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35836
A;Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77518; EMBL:Z23218
C;Superfamily: type E retrovirus env polyprotein

Query Match      88.7%; Score 173; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 4.8e-17;
Matches 31; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIIGDIRQAH 35
Db 26 CTRPNNTKSIHMGPGRAFYATGDIIGDIRQAH 60

RESULT 13
S35835
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35835
R;Chiodi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35835
A;Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77517; EMBL:Z23217
C;Superfamily: type E retrovirus env polyprotein

Query Match      88.7%; Score 173; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 4.8e-17;
Matches 31; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIIGDIRQAH 35
Db 26 CTRPNNTKSIHMGPGRAFYATGDIIGDIRQAH 60
```

QY 1 CSRPNNTKSIIPMGPGRAFYTGTGIIIGDIRQAH 35  
|.|||||  
Db 26 CTRPNNTKSIHMGPGRAFYTGTGIIIGDIRQAH 60  
|.|||||

## RESULT 14

S35848

envelope protein - human immunodeficiency virus type 1 (fragment)

C;Species: human immunodeficiency virus type 1, HIV-1

C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004

C;Accession: S35848

R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800

A;Accession: S35848

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-77 &lt;CHI&gt;

A;Cross-references: UNIPROT:Q77530; EMBL:Z23232; NID:G313607; PIDN:CAA80750.1; PID:G3136

C;Superfamily: type E retrovirus env polyprotein

Query Match 88.2%; Score 172; DB 2; Length 77;

Best Local Similarity 85.7%; Pred. No. 6.6e-17;

Matches 30; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTGTGIIIGDIRQAH 35  
|.|||||  
Db 26 CTRPNNTKSIHMGPGRAFYTGTGIIIGDIRQAH 60  
|.|||||

## RESULT 15

S35849

envelope protein - human immunodeficiency virus type 1 (fragment)

C;Species: human immunodeficiency virus type 1, HIV-1

C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004

C;Accession: S35849

R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800

A;Accession: S35849

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-77 &lt;CHI&gt;

A;Cross-references: UNIPROT:Q77531; EMBL:Z23233; NID:G313609; PIDN:CAA80751.1; PID:G3136

C;Superfamily: type E retrovirus env polyprotein

Query Match 88.2%; Score 172; DB 2; Length 77;

Best Local Similarity 85.7%; Pred. No. 6.6e-17;

Matches 30; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTGTGIIIGDIRQAH 35  
|.|||||  
Db 26 CTRPNNTKSIHMGPGRAFYTGTGIIIGDIRQAH 60  
|.|||||

Search completed: April 11, 2005, 11:15:28

Job time : 2.65208 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	195	100.0	866	2	Q9WP24	human
2	192	98.5	35	2	Q78198	human
3	192	98.5	91	2	O39362	human
4	192	98.5	91	2	O39692	human
5	189	96.9	85	2	O60956	human
6	189	96.9	88	2	O9DXK1	human
7	189	96.9	171	2	Q6UA59	human
8	188	96.4	35	2	O91430	human
9	188	96.4	35	2	O76PL9	human
10	188	96.4	35	2	Q76FM0	human
11	188	96.4	35	2	Q76FM1	human
12	188	96.4	35	2	Q76FM2	human
13	188	96.4	35	2	Q76FM3	human
14	188	96.4	35	2	Q76PM4	human
15	188	96.4	35	2	Q76PM5	human
16	188	96.4	35	2	Q76FM6	human
17	188	96.4	35	2	Q76FM7	human
18	188	96.4	35	2	Q76FM8	human
19	188	96.4	35	2	Q76PN0	human
20	188	96.4	35	2	Q76PN1	human
21	188	96.4	35	2	Q76PN2	human
22	188	96.4	35	2	Q76PN3	human
23	188	96.4	35	2	Q76PN4	human
24	188	96.4	35	2	Q77931	human
25	188	96.4	35	2	Q78039	human
26	188	96.4	35	2	Q787P3	human
27	188	96.4	35	2	Q787P4	human
28	188	96.4	35	2	Q787P5	human
29	188	96.4	35	2	Q787P6	human
30	188	96.4	35	2	O9QC04	human
31	188	96.4	92	2	O90205	human

GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Epidermis;  
 RX MEDLINE=94309197; PubMed=7545939;  
 RA Sala M., Zamburo G., Vartanian J., Marconi A., Bertazzoni U.,  
 RA Main-Hobson S.;  
 RT "Spatial discontinuities in human immunodeficiency virus type 1  
 RT quasiespecies derived from epidermal Langerhans cells of a patient with  
 RT AIDS and evidence for double infection.";  
 RL J. Virol. 68:5280-5283 (1994).  
 DR EMBL; Z34376; CAA84146.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR011056; Pept\_S24\_S26\_C.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 35 35  
 SQ SEQUENCE 35 AA; 3875 MW; 9F5A5B2A698C9ABE CRC64;  
 Query Match 98.5%; Score 192; DB 2; Length 35;  
 Best Local Similarity 97.1%; Pred. No. 3.5e-20;  
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35  
 Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35  
 RESULT 3  
 ID C039362 PRELIMINARY; PRT; 91 AA.  
 AC C039362;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98242906; PubMed=9583604;  
 RA Lukashov V.V., Op de Coul E.L., Coutinho R.A., Goudsmit J.;  
 RA "HIV-1 strains specific for Dutch injecting drug users in  
 RT heterosexually infected individuals in The Netherlands.";  
 RL AIDS 12:635-641 (1998).  
 DR EMBL; AF032157; AAC59382.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 91 91  
 SQ SEQUENCE 91 AA; 10363 MW; 0D9690FA2767592E CRC64;  
 Query Match 98.5%; Score 192; DB 2; Length 91;  
 Best Local Similarity 97.1%; Pred. No. 9.9e-20;  
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35  
 Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

Db 27 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 61  
 RESULT 4  
 ID Q69692 PRELIMINARY; PRT; 91 AA.  
 AC Q69692;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Envelope glycoprotein gp120 (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R923572;  
 RX MEDLINE=96392168; PubMed=8798980;  
 RA Lukashov V.V., Kuiken C.L., Boer K., Goudsmit J.;  
 RT "HIV type 1 subtypes in The Netherlands circulating among women  
 RT originating from AIDS-endemic regions.";  
 RL AIDS Res. Hum. Retroviruses 12:951-953 (1996).  
 DR EMBL; L76886; AAC37946.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 91 91  
 SQ SEQUENCE 91 AA; 10333 MW; 17273BE096D7593E CRC64;  
 Query Match 98.5%; Score 192; DB 2; Length 91;  
 Best Local Similarity 97.1%; Pred. No. 9.9e-20;  
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35  
 Db 27 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 61  
 RESULT 5  
 ID O90956 PRELIMINARY; PRT; 85 AA.  
 AC O90956;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Gp120 (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Heyndrickx L., Janssens W., Coppens S., Vereecken K., Willems B.,  
 RA Franssen K., Colebunders R., Vandenbruaens M., Van der Groen G.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ228189; CAA12819.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Glycoprotein; Transmembrane.  
 KW AIDS; Coat protein; Glycoprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 85 85  
 SQ SEQUENCE 85 AA; 9689 MW; 01BB36F634EA0AFB CRC64;

Query Match 96.9%; Score 189; DB 2; Length 85;  
Best Local Similarity 94.3%; Pred. No. 2.5e-19;  
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPWPGRAFYTGTGQIIGDIRQAH 35  
29 CTRPNNTKRSIPWPGKAFYTGTGQIIGDIRQAH 63

Db

RESULT 6

Q9DXK1 PRELIMINARY; PRT; 88 AA.

AC Q9DXK1; DB 2; Length 85;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN Name-env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP MEDLINE=21084776; PubMed=11216936;  
RX RA Op de Coul E.L.M., Prins M., Cornelissen M., van der Schoot A.,  
RA Boufassa F., Brettle R.P., Hernandez-Aguado L., Schiffer V.,  
RA McMenamin J., Rezza G., Robertson R., Zangerle R., Goudsmit J.,  
RA Coutinho R.A., Lukashov V.V.;  
RT "Using phylogenetic analysis to trace HIV-1 migration among western  
RT European injecting drug users seroconverting from 1984 to 1997.";  
RL AIDS 15:257-266 (2001).  
DR EMBL; AF307295; AAG32463.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000777; GP120.  
DR DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 88 88  
SQ SEQUENCE 88 AA; 9925 MW; DD523F6D428DB9E4 CRC64;

Query Match 96.9%; Score 189; DB 2; Length 88;  
Best Local Similarity 94.3%; Pred. No. 2.6e-19;  
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPWPGRAFYTGTGQIIGDIRQAH 35  
27 CTRPNNTKRSIPWPGKAFYTGTGQIIGDIRQAH 61

Db

RESULT 7

Q6UA59 PRELIMINARY; PRT; 171 AA.

AC Q6UA59;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN Name-env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang X.-P., Kahn P.D., Cunha C.B., Dosik D.A., Gawel C., Kaplan M.H.;  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AV375253; AA083712.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 171 171  
SQ SEQUENCE 171 AA; 18865 MW; 46C82EDEFBADE90B CRC64;

Query Match 96.9%; Score 189; DB 2; Length 171;  
Best Local Similarity 94.3%; Pred. No. 5.3e-19;  
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPWPGRAFYTGTGQIIGDIRQAH 35  
85 CSRPNNTKRSIPWPGKAFYTGTGQIIGDIRQAH 119

Db

RESULT 8

O91430 PRELIMINARY; PRT; 35 AA.

AC O91430;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN Name-env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98090117; PubMed=9430252;  
RA Ida S., Gatanaga H., Shioda T., Nagai Y., Kobayashi N., Shimada K.,  
RA Kimura S., Iwamoto A., Oka S.;  
RT "HIV type 1 V3 variation dynamics in vivo: long-term persistence of  
RT non-syncytium-inducing genotypes and transient presence of syncytium-  
RT inducing genotypes during the course of progressive AIDS.";  
RL AIDS Res. Hum. Retroviruses 13:1597-1609 (1997).  
DR EMBL; AB005412; BAA3329.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000777; GP120.  
DR InterPro; IPR011056; Pept\_S24\_S26\_C.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 35 35  
SQ SEQUENCE 35 AA; 3827 MW; 9F47EDF1198BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;  
Best Local Similarity 94.3%; Pred. No. 1.3e-19;  
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPWPGRAFYTGTGQIIGDIRQAH 35  
1 CARPNNTKRSIPGPGRAFYTGTGQIIGDIRQAH 35

Db

RESULT 9

Q76PL9 PRELIMINARY; PRT; 35 AA.

AC Q76PL9;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Envelope protein (Fragment).  
GN Name-env;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.

```
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34489; CAA84253.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

RESULT 10
Q76PM0 PRELIMINARY; PRT; 35 AA.
AC Q76PM0;
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34487; CAA84253.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

RESULT 11
Q76PM1 PRELIMINARY; PRT; 35 AA.
AC Q76PM1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Envelope protein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34487; CAA84253.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
```

```
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Envelope protein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34437; CAA84207.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

RESULT 12
Q76PM2 PRELIMINARY; PRT; 35 AA.
AC Q76PM2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Envelope protein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34435; CAA84205.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
```

```

FT  NON_TER      1
FT  NON_TER      35
SQ  SEQUENCE      35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match      96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps

QY  1 CSRPNNTKSPIMGPGRAFYTTGQIIGDIRQAHC 35
    1:|||||:|||||:|||||:|||||:|||||
Db   1 CTRPNNTKSPIGPGRAFYTTGQIIGDIRQAHC 35

RESULT 15
Q76PM5          PRELIMINARY;      PRT;      35 AA.
ID  Q76PM5
AC  Q76PM5;
DT  05-JUL-2004 (TREMBLrel. 27, Created)
DT  05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE  Envelope protein (Fragment).
GN  Name=env;
GC  Human immunodeficiency virus 1.
OC  Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX  NCBI_Taxid=11676;
[1]
RN  SEQUENCE FROM N.A.
RP  TISSUE=Epidermis;
RC  MEDLINE=94309197; PubMed=7545939;
RX  Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA  Wain-Hobson S.;
RT  "Spatial discontinuities in human immunodeficiency virus type 1
RT  quasiespecies derived from epidermal Langerhans cells of a patient with
RT  AIDS and evidence for double infection.";
RL  J. Virol. 68:5280-5283 (1994).
DR  EMBL; Z34430; CA84200.1; -.
DR  GO; GO:0019031; C:viral envelope; IEA.
DR  InterPro; IPR000777; GP120.
DR  InterPro; IPR011056; Pept_S24_S26_C.
DR  Pfam; PF00516; GP120; 1.
KW  AIDs; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT  NON_TER      1
FT  NON_TER      35
SQ  SEQUENCE      35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match      96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps

QY  1 CSRPNNTKSPIMGPGRAFYTTGQIIGDIRQAHC 35
    1:|||||:|||||:|||||:|||||:|||||
Db   1 CTRPNNTKSPIGPGRAFYTTGQIIGDIRQAHC 35

Search completed: April 11, 2005, 11:14:05
Job time : 12.206 secs

```

**THIS PAGE BLANK (USPTO)**